

17409
SEARCH REQUEST FORM

Examiner # (Mandatory): _____ Requester's Full Name: _____

Art Unit _____ Location (Bldg/Room#): 7E09 Phone (circle 305 306 308) _____

Serial Number: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

Title of Invention _____

Inventors (please provide full names): _____

Earliest Priority Date: _____

Keywords (include any known synonyms registry numbers, explanation of initialisms): _____

Search Topic:

Please write detailed statement of the search topic, and the concept of the invention. Describe as specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. You may include a copy of the abstract and the broadcast or most relevant claim(s).

STAFF USE ONLY

Searcher: PROB

Searcher Phone #: _____

Searcher Location: _____

Date Picked Up: 8-3Date Completed: 8-4-99Clerical Prep Time: 6 9Terminal Time: 6

Number of Databases: _____







Type of Search

☒ N.A. Sequence☒ A.A. Sequence☐ Structure (#)☐ Bibliographic☐ Litigation1☐ Fulltext☐ Procurement☐ Other

Vendors (include cost where applicable)

☐ STN☐ Questel/Orbit☐ Lexis/Nexis☐ WWW/Internet☒ MP3 In-house sequence systems (list)☐ Dialog☐ Dr. Link☐ Westlaw☐ Other (specify)

This Page Blank (uspto)

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn	n.a.	-	n.a.	database search, using Smith-Waterman algorithm

```
Run on: Tue Aug 3 21:17:35 1999; MasPar time 3626.14 Seconds
1536 403 Million cell updates/sec
```

Tabular output not generated.

Title:	>US-09-240-410-1
Description:	(1-2010) from US09240410.seq

N.A. Sequence: 1 ATGACGCTCTCCGCCCG.....TGCTCACTAGGCGCTCCG 2010

Comp:

Scoring table: TABLE default

```
Nmatch      STD : Dbase 0; Query 0
```

Searched: 646147 seqs, 1385953633 bases x 2

Post-processing:	Minimum Match	08
------------------	---------------	----

Database:

```

1:em_ba1 2:em_ba2 3:em_fun 4:em_htg 5:em_hun1 6:em_hun2
7:em_in 8:em_om 9:em_or 10:em_ov 11:em_pat 12:em_ph
13:em_pl 14:em_ro 15:em_sts 16:em_vi
genbank111
Database:

```

Database:

```

17:gb_ba1 18:gb_ba2 19:gb_htg1 20:gb_htg2 21:gb_in1
22:gb_in2 23:gb_com 24:gb_ov 25:gb_pat 26:gb_ph 27:gb_p11
28:gb_p12 29:gb_pr1 30:gb_pr2 31:gb_pr3 32:gb_ro
33:gb_st 34:gb_sts 35:gb_sy 36:gb_un 37:gb_vi

```

Statistics: Mean 11.468; Variance 5.912; scale 1.940

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2010	100.0	2594	31	AF071542	Homo sapiens semaphori	0.00e+00
2	2010	100.0	2661	31	AF069493	Homo sapiens GPI-anch	0.00e+00
3	2004	99.7	2656	31	AF030658	Homo sapiens semaphori	0.00e+00
4	1444	71.8	2608	32	AB017552	Mus musculus mRNA for	0.00e+00
5	847	42.1	1192	32	AF030659	Mus musculus semaphori	0.00e+00
6	375	18.7	8861	31	AF030657	Homo sapiens semaphori	9.69e-21
7	131	6.5	11874	37	AF018823	Alcelaphine herpesvtru	7.14e-78
8	131	6.5	130608	37	AF005370	Alcelaphine herpesvtru	7.14e-78
9	42	2.1	7218	25	166494	Sequence 14 from paten	2.65e-05
10	38	1.9	965	25	AR024229	Sequence 12 from paten	7.67e-07
11	38	1.9	1056	23	MV8H7256	Musclea vison GT dinuc	7.67e-07
12	38	1.9	2582	32	MUSIGCFRA	Mus musculus insulin-1	7.67e-07
13	38	1.9	7218	25	166494	Sequence 14 from paten	7.67e-07

C	15	38	1.9	8810	32	RNU59809	Rattus norvegicus	man	7.6e-07
	14	37	1.8	215	25	I28278	Sequence 5 from patent	3.06e-06	
	16	37	1.8	215	25	I28278	Sequence 5 from patent	3.06e-06	
C	17	36	1.8	965	25	AR024429	Sequence 22 from patent	1.20e-05	
	18	36	1.8	1056	23	MU087256	Mustela vison GT dinuc	1.20e-05	
C	19	37	1.8	1663	23	AF092534	Mustela vison microsat	3.06e-06	
	20	37	1.8	2336	24	AF022946	Gallus gallus microsat	1.20e-05	
	21	36	1.8	2919	31	HSU28369	Homo sapiens semaphori	1.20e-05	
C	22	36	1.8	8877	32	MM004710	Mus musculus domestica	1.20e-05	
	23	36	1.8	36534	30	U73167	Homo sapiens cosmid cl	1.20e-05	
C	24	36	1.8	36552	30	AF092098	Human chromosome 3p21	1.20e-05	
	25	36	1.8	123612	31	AC005024	Homo sapiens 3p21.1p1	1.20e-05	
C	26	35	1.7	74371	31	AC005369	Homo sapiens chromosome	4.56e-05	
	27	35	1.7	74371	31	AC005369	Homo sapiens chromosome	4.56e-05	
C	28	32	1.6	201	25	A10158	S. griseus gene for pre	2.43e-03	
	29	32	1.6	201	25	A10158	Synthetic DNA for prep	2.43e-03	
C	30	32	1.6	201	25	A10161	S. griseus gene for prep	2.43e-03	
	31	32	1.6	461	32	MUSINSUL01	Mouse insulin-like gro	2.43e-03	
C	32	32	1.6	879	30	HSU29899	Homo sapiens clone Drl	2.43e-03	
	33	32	1.6	1270	32	RATMGAA21	Rat alpha-2-macroglobu	2.43e-03	
C	34	32	1.6	1864	30	AF020276	Homo sapiens spinocere	2.43e-03	
	35	32	1.6	2040	23	CFU29765	Cantis familiaris alpha	2.43e-03	
C	36	32	1.6	2265	32	AF080091	Mus musculus semaphori	6.62e-04	
	37	33	1.6	2383	32	AF080090	Mus musculus semaphori	6.62e-04	
C	38	32	1.6	3367	31	AF032105	Homo sapiens ataxin-7	2.43e-03	
	39	33	1.6	3518	32	RATN1TB	Rattus norvegicus neur	6.62e-04	
C	40	32	1.6	3819	31	AF106861	Homo sapiens attractin	2.43e-03	
	41	32	1.6	3969	30	HS5CA7	Homo sapiens mRNA for	2.43e-03	
C	42	32	1.6	4029	32	MM026348	Mus musculus insulin-1	2.43e-03	
	43	32	1.6	5633	29	HSU17033	Human 180 kDa transmem	2.43e-03	
C	44	33	1.6	7913	23	BOVCMRP	Bovine cation-independ	6.62e-04	
	45	32	1.6	8670	27	CRAKRG7	Chlamydomonas reinhard	2.43e-03	

ALIGNMENTS

*
JOURNAL MEDLINE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK COMMENT
FEATURES
SOURCE

CDS

1 AF071542 2594 bp mRNA PRI 01-DEC-1998
Homo sapiens semaphorin K1 mRNA, complete cds.
AF071542
g3930578
AF071542.1 GI:3930578
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 2594)
Xu,X., Ng,S., Wu,Z.L., Nguyen,D., Homburger,S., Seidel-Dugan,C.,
Ebbers,A. and Luo,Y.
Human semaphorin K1 is glycosylphosphatidylinositol-linked and
defines a new subfamily of viral-related semaphorins
J. Biol. Chem. 273 (35), 22428-22434 (1998)
98380463
2 (bases 1 to 2594) August 78 include motif
Xu,X. and Luo,Y.
Direct Submission
Submitted (10-JUN-1998) Biology, Exelixis Pharmaceuticals, Inc., 2606
Littlefield Ave, South San Francisco, CA 94080, USA
3 (bases 1 to 2594)
Xu,X. and Luo,Y.
Direct Submission
Submitted (30-NOV-1998) Biology, Exelixis Pharmaceuticals, Inc., 2606
Littlefield Ave, South San Francisco, CA 94080, USA
Sequence update by submitter
On Dec 1, 1998 this sequence version replaced gi:3493630.
Location/Qualifiers
1..2594
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q22"
1..2001

RESULT	2	LOCUS	AF069493	2661 bp	mRNA	PRI	03-DEC-1998
DEFINITION	Homo sapiens GPI-anchored membrane protein CDw108 precursor, mRNA, complete cds.						
ACCESSION	AF069493						
NID	93947939						
VERSION	AF069493.1 GI:3947939						
KEYWORDS	human,						
SOURCE	Homo sapiens						
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE	1 (bases 1 to 2661)						
AUTHORS	Yamada, A., Kubo, K., Takeshita, T., Harashima, N., Kawano, K., Segawa, K., Sugamura, K. and Itoh, K.						
TITLE	Direct Submission						
JOURNAL	Submitted (02-JUN-1998) Immunology, Kurume University School of Medicine, Asahi-machi 67, Kurume, Fukuoka 830, Japan						
FEATURES	<p>location/Qualifiers</p> <p>1..2661</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/chromosome="15"</p> <p>/map="between WI-6247 and CHLC.GARA5D02"</p> <p>/cell_line="HPB-ALL"</p> <p>/cell_type="T lymphocyte"</p> <p>18..2018</p> <p>/note="putative adhesion molecule; lymphocyte activation antigen"</p> <p>/codon_start=1</p> <p>/product="GPI-anchored membrane protein CDw108 precursor</p> <p>/protein_id="AAC82642.1"</p> <p>/db_xref="PID:93947940"</p> <p>/db_xref="GI:3947940"</p> <p>/translation="MPPPPGARAAPSAPRARVGPPTLGLPLRLRLLLLLMAAAAG OGLRSGPRIPFAWMKGVGDGDDVDGDEGDEPTVLFHEPSSSWVGGRGVYLFDFP GKAASRTVNIIGSTKSCDLKRDCENYITLERSBGLIACGTAAHPSCNLYNGI VPLGEGRGAPSPDENSLVFEQGEVSTLRKDEYNKIRFRRIGSELYTSIDT MNPQDIFKATIHODQAYDKIYFPERDNDKNEAPLNIRVAOLCRGOGGESS SVSKWNTFLKAMLVCSDAATNFKNRLLDVFLLPSSQMDTNYGVFSNPNYS CYSLLIDIDIKVFTSLKIGHSLPLNPPRGCIIPQOQIPIETFOVADRHEVOR PMKLTPLFHSKHYHOKYAVHRMOASGEPHYLYLTDRGTHKYVVEGDEGHS FNIMELOPFRRRAAIOTMKSIDERRKLYVSSQVEYSPDLDCYTVGGGCHGLMS PYCGWQDGRGISIYSSERSVLQSIINPAPPHKECPPKPKDKAPLQKSLAPASRYLS PMSRRAITSMRKNEVQSCPEGHQSNCLIFENLTAQOYGYHPCQEADEGSYFR OHMQLPEDGIMAEHLHAGACLAASLWGLVPTLTGLIVH"</p>						
DB	1861	TTCCGAGAGGCTCAGACACTGGACAGCTCTGCCCCGAGAGCGCATCATGGCCGAGCACTG	1920				
OY	1861	TTCCGAGAGGCTCAGACACTGGACAGCTCTGCCCCGAGAGCGCATCATGGCCGAGCACTG	1920				
DB	1921	CTGGGTCATGCTCTTGCCCTTGCCCTGCTCCCTTGCGCTGGGGGTGCTGCCCACTACT	1980				
OY	1921	CTGGGTCATGCTCTTGCCCTTGCCCTGCTCCCTTGCGCTGGGGGTGCTGCCCACTACT	1980				
DB	1981	CTTGCGTTGCTGTCACACTAGGGGCTCCCG	2010				
OY	1981	CTTGCGTTGCTGTCACACTAGGGGCTCCCG	2010				
misc_feature	/note="glycosylation site"						
misc_feature	/evidence="not-experimental"						
misc_feature	/note="glycosylation site"						

[illegible]

REFERENCE
1 (sites)
AUTHORS
TITLE
Molecular cloning and expression of novel Semaphorin family of molecule
JOURNAL
Unpublished (1998)
2 (bases 1 to 2608)
AUTHORS
TITLE
Takahashi, H. and Sato, Y.
Direct Submission
Submitted (08-SEP-1999) to the DDBJ/EMBL/GenBank databases. Hiroshi Takahashi, Mitsubishi Kasei Institute of Life Sciences, Neurobiology Lab; 11 Minamiooya, Machida, Tokyo 194, Japan (E-mail: hiroshi@libra.is.m.kagaku.co.jp, Tel: 81-427-24-6211, Fax: 81-427-24-6314)

FEATURES
source
Location/Qualifiers

1..2608
/organism="Mus musculus"
/db_xref="taxon:10090"
/dev_stage="postnatal"
/tissue_type="brain"
5..2000
/gene="msemk1"
6..2000
/gene="msemk1"
/codon_start=1
/product="msemk1p"
/protein_id="BAA75665.1"
/db_xref="PID:d1038405"
/db_xref="GI:4519588"
/translation="MTPPPGRAAPSAPRARVLSLPARFGLRLRLILFVWVAASA
OCHSRSGPRISAVMKGDHVDSEPEPHVLFHEGSEFVWVGSGKXVHEFPEGN
ASRVYINIGSTGSCODKODGNITLLERNGNLVGTNARKSCWMLVNDSTVMS
LGEWGAPSPDENSTYLFEGDEVYSTRKDEYNGKIPFRFRIGESLITSDTVMQ
NPQFIKATIVHDDQAYDDKITFFREDNEDKNEPPLANRYAOLCRGQGGESSLSV
SKWNTFLKAMLVSDAATNRNRLQDVFLLDPESQWMDIVYGVFSNPMYSAVCV
YSLGIDIVERTSLKGYHMG.LPNRPQGLPKQPIETFOVADSHVEVNAQREPM
GLKPLTFHSKYHOKVYVYRMQASNGEFLVLTLDGTLHKVSGSDGDSHFVN
IMEIOPFHRAAIOAISLDADRKLVTYSOMESVPLDMCEVYSGGCLMSDPMY
CGWDDRCVSTISOSRVSLOSTINPAEPHECNPBPDEAPLOKYSIARSRYYLPCP
ESRNATYILMRHENVBOSECPGHOSPCSLFIENITARYGHTKCEADQBSTLREAOH
WELLPEDRALQIMHNGARALASFWLGVLPFLILGLLVH"

BASE COUNT 554 a 769 c 727 g 558 t
ORIGIN

Query Match 71.8% Score 1444: DB 32: Length 2608:
Best Local Similarity 87.4% Pred. No. 0.00e+00: Mismatches 241: Indels 12: Gaps 2:
Matches 1757: Conservative 0:

Db 6 ATGACGCTCTCTCCCGGACGTGCGGCCCGGACGACCGCGCGCGCTCTCAGC 65
Oy 1 ATGACGCTCTCTCCCGGACGTGCGGCCCGGACGACCGCGCGCGCTCTCAGC 60
Db 66 CTGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 125
Oy 61 CCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 126 GCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 184
Oy 121 GCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 185 -----GCAGACCATGTGACCTTATGACCATGTGACCATGTGACCATGTGACCAT 236
Oy 181 CATGTAGGACGACGCGGTGACCTTGTGACCATGTGACCATGTGACCATGTGACCAT 240
Db 237 GAGCGGAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 296
Oy 241 GAGCGGAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 297 CCGGAGGAGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 356
Oy 301 CCGGAGGAGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360

Db 357 CAGACAAACAGGACCTGTGGAAATTACATCAGCTTCTGAAAGCGGGGTATGGGCTG 416
Oy 361 CTGGATTAAGCGGACCTCCGAAGTACATCTCTCTGGAAGGGGGATGAGGGGCTG 420
Db 417 CTGGCTGTGGACCAATGCCCGGAAGCCAGCTGTGAACTTGTGATGACAGTGTG 476
Oy 421 CTGGCTGTGGACCAAGCCCGGACCCAGCTGTGAACTTGTGATGAGGACTGTG 480
Db 477 GTATGTCACTGTGTAGATGAAGGCTATGCCCTTTAGCCCGGATGAGAACTTCTG 536
Oy 481 GTGC---CACTTGTGCAATGAGAGGCTACGCCCTTACGCCGACGAGAACTTCTG 537
Db 537 GTTCTGTTGAAGAGATGAAGGTCTCTTACATCCGGAACGAGATACAAACGGGAAG 596
Oy 538 GTTCTGTTGAAGAGAGATGAAGGTCTTACATCCGGAACGAGATACAAACGGGAAG 597
Db 597 ATCCCTCGGTTTGCAGCATTCGGGGCGAGGTGAGTGAAGTACACAAGTATCAGTATG 656
Oy 598 ATCCCTCGGTTTGCAGCATTCGGGGCGAGGTGAGTGAAGTATCAGTATCAGTATG 657
Db 657 CAGAACCCACAGTTCTATCAGGCCACCATTTGTGACCAAGACCAAGCTATGATGATAAG 716
Oy 658 CAGAACCCACAGTTCTATCAGGCCACCATTCGTGACCAAGACCAAGCTTACGATGACAAG 717
Db 717 ACTTACTTCTTCGCGAAGACCAACCTGACCAAGAACCCCGAGGCTCTCTCAATGTG 776
Oy 718 ACTTACTTCTTCGCGAAGACCAANTCTGCAAGAACTCTGAGGCTCTCTCAATGTG 777
Db 777 TCCCGATGAGCCAGTTGTGACGAGGGGGACCAAGGTGTGAGAGTTGTTGTCTCTCC 836
Oy 778 TCCCGATGAGCCAGTTGTGACGAGGGGGACCAAGGTGTGAGAGTTGTTGTCTCTCC 837
Db 837 AAGTGAACACTTCTCTGAAGACCATGTTGCTCTGACGATGACGACCAACAGCAAC 896
Oy 838 AAGTGAACACTTCTCTGAAGACCATGCTGATGATGATGATGATGATGATGATGATG 897
Db 897 TTCAATCGGCTCAAGATGTTCTCTGCTCCCTGACCCAGGAGGAGGAGGAGATATCC 956
Oy 898 TTCAACAGGCTCAAGAGCTCTCTGCTCTCTGACCCAGGAGGAGGAGGAGGAGGAG 957
Db 957 AAGGCTATGAGGCTTCTTCCAAACCCCTGGAACCTAGCTAGCTGTGCTGTATTCCTT 1016
Oy 958 AAGGCTATGAGGCTTCTTCCAAACCCCTGGAACCTAGCTAGCTGTGCTGTATTCCTT 1017
Db 1017 GGTGACATTGAGAGGCTCTCGTACCTCATGCTTCAAGAGGCTTACCATATGGGCTTCC 1076
Oy 1018 GGTGACATTGAGAGGCTCTCGTACCTCATGCTTCAAGAGGCTTACCATATGGGCTTCC 1077
Db 1077 AACCTCGACCTGGGATGCTCTCCCAAAAAAGAGCCATACCCAGAAACCTTCCAG 1136
Oy 1078 AACCTCGGCTGGGATGCTCTCCCAAGAGGCTCTCCCAAGAGGCTTACCATATGGGCTTCCAG 1137
Db 1137 GTAGCTGATAGTACACCAAGAGGTGCTCAGAGGGTGAAGCTATGAGGCGCACTGAAGACA 1196
Oy 1138 GTAGCTGATAGTACACCAAGAGGTGCTCAGAGGGTGAAGCTATGAGGCGCACTGAAGACA 1197
Db 1197 CCATTGTTCCATTCTTAAGTATACCATTAACCAAGAGGTGCTTACCCGATGCAAGCAGC 1256
Oy 1198 CCATTGTTCCATTCTTAAGTATACCATTAACCAAGAGGTGCTTACCCGATGCAAGCAGC 1257
Db 1257 AATGAGAGACCTTCCATGCTGTTATATCAACAGAGAGAGGAGGACCATTCACAAAGGTG 1316
Oy 1258 CAGGAGGAGACCTTCCATGCTGTTATATCAACAGAGAGAGGAGGACCATTCACAAAGGTG 1317
Db 1317 GTGAAATCAGGAGGAGACAGGACCATGCTTGTCTCAATATATCATGAGATCAGGCTTCT 1376
Oy 1318 GTGAAATCAGGAGGAGACAGGACCATGCTTGTCTCAATATATCATGAGATCAGGCTTCT 1377
Db 1377 CACGCTGAGAGGAGGACATCAGGCTATATCATGAGTGTGACCGGAGGAGGAGGAGGAG 1436
Oy 1378 CACGCTGAGAGGAGGACATCAGGCTATATCATGAGTGTGAGGAGGAGGAGGAGGAGGAG 1437
Db 1437 ACCTCCAGTGAAGATGAGGACGAGTACCTGACATGTGTGAGGCTTACAGGAGGAGGAGC 1496

QY	1438	AGCTCCAGTGGGAGGTAGCCAGGTGGCCCTGGACGTGTGTGAGTGTATGCGGGGGC	1497
Db	1497	TGCCATGGCTGCTCATGTATCCCGAGACCCCTACTGTGGCTGGGACAGACCGCTGCGTG	1556
QY	1498	TGCCACGGTTCCTCATGTATCCCGAGACCCCTCTACTGCGCTGGACACAGGCGCGTGCATC	1557
Db	1557	TCTATCTACAGTCTCCCAAGGATGATGCTGCTGCAGTCCATTAACTCAAGCGGAGACACAGAA	1616
QY	1558	TCCATCTACAGTCTCCGAAGGGTCATGTGTGTCATCATTTAACTCAAGCGGAGACACAGAA	1617
Db	1617	GAGTGTCCCAACCCCTAAACACAGACAGAGGCTCCATTCAGAGAAAGTTCCCTGGCTCGAAC	1676
QY	1618	GAGTGTCCCAACCCCAACACAGACAGAGGCTCCATTCAGAGAAAGTTCCCTGGCCCAAC	1677
QY	1678	TCTGCTACTACTAGCTGAGCTGGCCCATGAGATCCCGCCACAGCCACTACTCATGAGCGCCAC	1737
Db	1737	GAGGAGATGTGTGACACAGAGCTGTAGCCAGGCGCCACCAAGCCCTAGTGCATCTCTGTTCC	1796
QY	1738	AAGGAGAACGTGTGACACAGAGCTGTAGCCAGGCGCCACCAAGCCCTAGTGCATCTCTGTTCC	1797
Db	1797	ATTGAGACCTACAGGGCCGCTCAGTATGAGCCACTACCGCTGGAGAGCCCAAGAGGCTCC	1856
QY	1798	ATCGAGAACCTACAGGGCCGCTCAGTATGAGCCACTACCTACTCTGGAGAGCCCAAGAGGCTCC	1857
Db	1857	TACCTCCGTGAGGCTCAACACTGTGGAGGCTGTGTCGCCAGAGACAGACACTGGCCGAGCAA	1916
QY	1858	TACTCCCGAGAGGCTCAGACACTGAGCAGCTGTGTCGCCAGAGACGAGCATATGCGCCAGCAC	1917
Db	1917	CTAATGGGCACTGCCCCGGGCTCTGGCCGCTCTCTGTGGCTGGGGGTCTCTGCCACACTC	1976
QY	1918	CTGCGGGCTCATGCTGCTGGCCCTGGCCCTCTCTCTGTGGGGGTGCTGCCACACTC	1977
Db	1977	ATACTGTGCTGTGCTGTCTCATCAGTGGGCTTC	2006
QY	1978	ACTCTGGCTGTGCTGTCTCATCAGTGGGCTTC	2007
RESULT	5		
LOCUS	AF030699	1192 bp	mRNA
DEFINITION	Mus musculus semaphorin L (Semal) mRNA, partial cds.	ROD	04-SEP-1998
ACCESSION	AF030699		
NID	93523116		
VERSION	AF030699.1	GI:3523116	
KEYWORDS			
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
AUTHORS	Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 1192)		
JOURNAL	Lange, C., Liehr, T., Goen, M., Gebhart, E., Fleckenstein, B. and		
LINE	Ensser, A.		
REFERENCE	New eukaryotic semaphorins with close homology to semaphorins of		
AUTHORS	DNA viruses		
TITLE	Genomics 51 (3), 340-350 (1998)		
JOURNAL	98389619		
LINE	2 (bases 1 to 1192)		
REFERENCE	Ensser, A., Liehr, T., Lange, C., Goen, M., Gebhart, E. and		
AUTHORS	Fleckenstein, B.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-OCT-1997) Institut fuer Klinische und Molekulare		
LINE	Virologie, Friedrich Alexander Universitaet, Schlossgarten 4,		
REFERENCE	Erlangen, Bavaria 91054, Germany		
AUTHORS			
FEATURES	Location/Qualifiers		
SOURCE	1..1192		
gene	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/map="9A3..3-B"		
	/chromosome="9"		
	1..>1192		
	/gene="Semal"		

CDS		/evidence=not_experimental 13.. >1192 /gene="SemaL" /note-"M-Sema-L" /codon_start=-1 /evidence=not_experimental /product="semaphorin L" /protein_id="AAC34262.1" /db_xref="PID:g3523117" /db_xref="GI:3523117"		
BASE COUNT	254 a	355 c	336 g	247 t
ORIGIN				
Query Match	42.1%;	Score 847;	DB 32;	Length 1192;
Best Local Similarity	88.1%;	Pred.	No. 0.0e+00;	
Matches 1048;	Conservative	0;	Mismatches 129;	Indels 12; Gaps 2.
Db	13	ATGAGCGCTCCTCCCTCCGGACGTGCCCCGCCCAAGGCACGGCGCGCCGGTCCTAGC	72	
Oy	1	ATGAGCGCTCCTCCGGACGTGCCCCGCCCAAGGCACGGCGCGCCGGTCCTAGC	60	
Db	73	CTGCGGCGCTCGGTTTCGGGACTCCCCTCGGCTCGGCGCTTGCTGCTGTGGTGGCC	132	
Oy	61	CCGCGCGCTCGGTTTCGGGACTCCCCTCGGCTCGGCGCTTGCTGCTGTGGTGGCC	120	
Db	133	GCCGCGCTCCGCCCAAGGCGACTCGAGAGCGGACC CGGCATCTCCGCGTCTGGAAAG-	191	
Oy	121	GCCGCGCTCCGCCCAAGGCGACTCGAGAGCGGACC CGGCATCTCTCGCGTCTGGAAAGC	180	
Db	192	-----GCAGAACATGTGGATTATTAAGCCAGCTGAGCCACACACCGCTTTTCAT	243	
Oy	181	CATGTAGGCGAGGACCGGGGTGACATTTGGCCAGCATGAGCGCACAGGCGCTTTTCAC	240	
Db	244	GAGCGGGGACACTTCTCTGTGTGGGTGGGTGAAGSTGGCAAGGTCTACCACTCAATTC	303	
Oy	241	GAGCGGGGACACTTCTCTGTGTGGGTGGGTGAAGSTGGCAAGGTCTACCTTTGACTTC	300	
Db	304	CCCGAGGGCAAGAAATGCCCTCTGTGCGACGCGTGAACAATCGGCTCCACAAAAGGGTCTGT	363	
Oy	301	CCCGAGGGCAAGAAATGCCCTCTGTGCGACGCGTGAATTCGGCTCCACAAAAGGGTCTGT	360	
Db	364	CAGGACAAACAGACTGTGGGAATTACATCTCTTTAGAAAGCGGGGTAATGGCTG	423	
Oy	361	CTGGATTAAGCGGAGCTGCGAAMATACTACTCTCTGTGAGAGCGCGAGTGAAGGGCTG	420	
Db	424	CTGGCTGTGGCACCAATGCCCGGAAGCCAGCGTGGGAATGGGTAATGACAGTGTG	483	
Oy	421	CTGGCTGTGGCACCAATGCCCGGAAGCCAGCGTGGGAATGGGTAATGACAGTGTG	480	
Db	484	GTGATGTACTTGTGTGAATGAAGGCTATCCCCCTTTCAGCCCGGATGGAATCCCTG	543	
Oy	481	GTGC---CACTTGGCGAATGAGAGGCTAGCCCCCTTTCAGCCCGGAGAGAATCCCTG	537	
Db	544	GTTCTGTTTGAAGAGATGAAGTACTTACATCCGGAAGCAGGAATACACAGGGAAG	603	
Oy	538	GTTCTGTTTGAAGAGAGATGAAGTACTTACATCCGGAAGCAGGAATACATAGGGAAG	597	
Db	604	ATCCCTGSGTTTCGACGATTCGGGGGGAAGTGAATGTAAACAATGATTAACGTCATG	663	
Oy	598	ATCCCTGSGTTTCGACGATTCGGGGGGAAGTGAATGTAAACAATGATTAACGTCATG	657	
Db	664	CAGAACCACAGTTCATCAAGGCGCACCATTTGTGACCAACAACCAAGCTATGATGAATAG	723	
Oy	658	CAGAACCACAGTTCATCAAGGCGCACCATTTGTGACCAACAACCAAGCTTACGATGAACAAG	717	

DB	724	ATCTACTACTTCTCCGAGAACAAACCTGACAGAACCCGAGGCTCCTTCATGtg	783
Oy	718	ATCTACTACTTCTCCGAGAGACACTCTGACCAAGAAATCTGAGGCTCCTTCATGtg	777
Db	784	TCCCGAGTgGCCAGTgTGCGAGGGGGGACGAGGCGTgGAGAGTTCGTGTCTCTCC	843
Oy	778	TCCCGTGTGGCCAGTGTGGCAGGGGGACGAGGCGTgGAGAGTTCGTGTCTCTCC	837
Db	844	AAGTGGACACCTTCTCTGGAAGCCATGTGTCTGTCGAGCATGACGCCACCAAGGAAC	903
Oy	838	AAGTGAACACATTTCTTGAAGCCATGTGTATGAGTgATGAGTGTGTGTCCACCAAGAAAC	897
Db	904	TTCAATCGGCTGCAAGATGTCTTCTGCTTCCTGACCCCAAGTGGCCAGTGAAGATAC	963
Oy	898	TTCAAGAGGCTGCACAAACCTCTTCTGCTTCCTGACCCCAAGTGGCCAGTGAAGATAC	957
Db	964	AGGGCTATAGGGGTTTTCTTCCAACCCCTGAGACTACTGATGTGTGTGTATTCGTT	1023
Oy	958	AGGGCTATAGTGTGTGTCTTCTTCCAACCCCTGAGACTACTGATGTGTGTATTCGTT	1017
Db	1024	GGTGACATTTAGACAGATCTTCCGTACTCTATGCTTCACAAAGGCTTACACATGGAGCTTCC	1083
Oy	1018	GGTGACATTTAGACAGATCTTCCGTACTCTATGCTTCACAAAGGCTTACACATGGAGCTTCC	1077
Db	1084	AACCTTCGACCTGGGATGTGCTCTCCCAAAAACGACCCCATATCCACAGAAACCTTCAG	1143
Oy	1078	AACCTTCGACCTGGGATGTGCTCTCCCAAAAACGACCCCATATCCACAGAAACCTTCAG	1137
Db	1144	GTAGCTGATAGTACCCAGAGAGTGGCTGACAGAGGAGGAGACCATATGGAGG	1192
Oy	1138	GTGGCTGACCGTACACCCAGAGAGTGGCTGACAGAGGAGGAGACCATATGGAGG	1186
RESULT	6		
LOCUS	AF030697	8861 bp	DNA
DEFINITION	Homo sapiens semaphorin L (SEMA1) gene, partial cds.		
ACCESSION	AF030697		
NTD	g3551778		
VERSION	AF030697.1	GI:3551778	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
AUTHORS	Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 8861)		
JOURNAL	Lange, C., Liehr, T., Goen, M., Gebhart, E., Fleckenstein, B. and		
MEDLINE	Emser, A.		
REFERENCE	New eukaryotic semaphorins with close homology to semaphorins of		
AUTHORS	DNA viruses		
TITLE	Genomics 51 (3), 340-350 (1998)		
JOURNAL	2. (bases 1 to 8861)		
MEDLINE	Emser, A., Liehr, T., Lange, C., Goen, M., Gebhart, E. and		
REFERENCE	Fleckenstein, B.		
AUTHORS	Direct Submission		
TITLE	Submitted (72-OCT-1997) Institut fuer Klinische und Molekulare		
JOURNAL	Virologie, Friedrich Alexander Universitaet, Schlossgarten 4,		
source	Erlangen, Bavaria 91054, Germany		
FEATURES	location/Qualifiers		
source	1..8861		
source	/organism="Homo sapiens"		
source	/db_xref="taxon:9606"		
source	/map="15p22.3-p23"		
source	/chromosome="15"		
source	join(1..114,605..646,944..1036,1238..1322,1443..1553,		
source	2176..2315,2907..3091,3945..4053,4146..4344,6876..7003)		
source	7178..7332,7533..7594,7904..>8265)		
source	/gene="SEMA1"		
source	/product="semaphorin L"		
source	/evidence=experimental		
source	<1..>8265		
source	/gene="SEMA1"		
source	join(1..114,605..646,944..1036,1238..1322,1443..1553,		
source	2176..2315,2907..3091,3945..4053,4146..4344,6876..7003)		
source	7178..7332,7533..7594,7904..>8265)		
source	/gene="SEMA1"		
source	/product="semaphorin L"		
source	/evidence=experimental		
source	<1..>8265		
source	/gene="SEMA1"		
source	join(1..114,605..646,944..1036,1238..1322,1443..1553,		
source	2176..2315,2907..3091,3945..4053,4146..4344,6876..7003)		
source	7178..7332,7533..7594,7904..>8265)		
source	/gene="SEMA1"		
source	/product="semaphorin L"		
source	/evidence=experimental		
source	<1..>8265		
source	/gene="SEMA1"		
source	join(1..114,605..646,944..1036,1238..1322,1443..1553,		
source	2176..2315,2907..3091,3945..4053,4146..4344,6876..7003)		
source	7178..7332,7533..7594,7904..>8265)		
source	/gene="SEMA1"		
source	/product="semaphorin L"		
source	/evidence=experimental		
source	<1..>8265		
source	/gene="SEMA1"		
source	join(1..114,605..646,944..1036,1238..1322,1443..1553,		
source	2176..2315,2907..3091,3945..4053,4146..4344,6876..7003)		
source	7178..7332,7533..7594,7904..>8265)		
source	/gene="SEMA1"		
source	/product="semaphorin L"		
source	/evidence=experimental		
source	<1..>8265		
source	/gene="SEMA1"		
source	join(1..114,605..646,944..1036,1238..1322,1443..1553,		
source	2176..2315,2907..3091,3945..4053,4146..4344,6876..7003)		
source	7178..7332,7533..75		

Query Match	18.7%	Score 375	DB 31	Length 8861
Best Local Similarity 100.0%		Pred. No. 9.69e-270		
Matches 375: Conservative	0	Mismatches 0	Indels 0	Gaps 0
Db 7900 CCAGACAAAGGCCCCACAGCGGAGAGGTTCCCTGGCCCCCAACACTCTGCTACTACTGAGC 7959				
Qy 1636 CCAGACAAAGGCCCCACAGCGGAGGTTCCCTGGCCCCCAACACTCTGCTACTACTGAGC 1695				
Db 7960 TGGCCCATGGAAATCCCGCCAGCCGACCTACTCATATGGGCCACAGAGGAAGCTGGAGCAG 8019				
Qy 1696 TGGCCCATGGAAATCCCGCCAGCCGACCTACTCATATGGGCCACAGAGGAAGCTGGAGCAG 1755				
Db 8020 AGCTGGCAACCTGTGTACCAAGAGCCCAACTGCATCTGTTTCATCGAAGAACTTCAGGCG 8079				
Qy 1756 AGCTGGCAACCTGTGTACCAAGAGCCCAACTGCATCTGTTTCATCGAAGAACTTCAGGCG 1815				
Db 8080 CAGCACTAAGGGCCACTACTTTTGGCGAGGCCAGAGAGGCTCCTACTTCGCGAGGCTCAG 8139				
Qy 1816 CAGCACTAAGGGCCACTACTTTTGGCGAGGCCAGAGAGGCTCCTACTTCGCGAGGCTCAG 1875				
Db 8140 CACTGCGACCTGCTGCGCCAGAGCGGATCATGAGCGGAGCACTGCTGGTCTATGCTGT 8199				
Qy 1876 CACTGCGACCTGCTGCGCCAGAGCGGATCATGAGCGGAGCACTGCTGGTCTATGCTGT 1935				
Db 8200 GCCCTGCGCCGCTCCCTCTGCGTGGGGGTGTGTCACACATCACTTGGCTGTGTCGTC 8259				
Qy 1936 GCCCTGCGCCGCTCCCTCTGCGTGGGGGTGTGTCACACATCACTTGGCTGTGTCGTC 1995				
Db 8260 CACTAGGGGCTCCCG 8274				
Qy 1996 CACTAGGGGCTCCCG 2010				

BASE COUNT 1749 a 2358 c 2757 g 1997 t

ORIGIN

2176 .2315,2907. .3091,3945. .4053,4146. .4344,6876. .7003

7178. .7332,7533. .7594,7904. .8265)

/gene="SEMA1"

/note="sequence starts with the putative second exon of

SEMA1: H-SEMA-L"

/codon_start=1

/product="semaphorin L"

/protein_id="AAC34741.1"

/db_xref="PID:g3551779"

/db_xref="GI:3551779"

/translation="EPHTVLFHEPGRSSVWVGGRGVYLLDFPEGRNASRYTNIGST

KSGCLRDPCENYITTLERSSELLCGSTRARPSGMLYNTGVPLGEMRGTAFTSP

DENSLVLEBDEYSTRKOEINGKIPRRIRIGESLELYTSDVMONPOFKATIVHQ

DAVDKIKYIFEREDPNLNPENAPLQSRDQDGESSLYSKMTFLKAMLY

CSAAATKNFNRLQDVELLPDPSGWRDTRVYVFENPMYSAVCYSGLDDIDKVRT

SLKGYHSSILPNRPGRKCLPDQDPIETFEFQVADRPEVAORPEMGPILKTPLEHRY

HSQVAVHMOASHGTFPHVLYLTTRGTLHKVYEGEDQESHAFNIMELQPRRAA

IQTMSIDAERRKILYVSQWESQVPLDLELVGGGHGCLMSDPYCGWQGRICISY

SSRSYLOSINPAPHEKPCPNKAPLQKXSLANPSRYLSCPMSRATYSWRHK

ENVEQCEBQHSQPCILFIENLTAAQVGYHCEAOGESYFREAOHWMLPEGIMAE

HLGHACALAAASIMLVGLVPLTFLGLLVH"

RESULT 7

LOCUS AHU18243 11874 bp DNA

DEFINITION Alcelaphine herpesvirus 1 putative semaphorin homolog (AHV-sema)

and putative membrane antigen genes, complete cds, and major

sDNA-binding protein gene, partial cds.

ACCESSION U18243

NTD g1000716

VERSION U18243.1 GI:1000716

KEYWORDS

SOURCE Alcelaphine herpesvirus 1.

ORGANISM Alcelaphine herpesvirus 1.

REFERENCE 1 (bases 1 to 11874)

Enser, A. and Fleckenstein, B.

Alcelaphine herpesvirus type 1 has a semaphorin-like gene

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

Gammaherpesvirinae; unclassified Gammaherpesvirinae.

JOURNAL	J Gen. Virol. 76 (Pt 4), 1063-1067 (1995)
MEDLINE	97201573
REFERENCE	2 (bases 1 to 11874)
AUTHORS	Ensser, A.
TITLE	Direct Submision
JOURNAL	Submitted (07-DEC-1994) Armin Ensser, Institut fuer Virologie, Friedrich-Alexander-Universitaet Erlangen-Nuernberg, Schlossgarten 4, Nuernberg, Bavaria 91054, Germany
source	Location/Qualifiers 1. 11874
	/organism="Alcelaphine herpesvirus 1"
	/strain="C500"
	/specific_host="Carnophantes taurinus taurinus"
	/db_xref="taxon:35252"
	/lab_host="Bos taurus"
	/clone="clones C238, C249, C500BS"
	<1. (330, 346)
repeat_region	/note="gamma-2 herpesviral GC-rich terminal tandem repeats"
	/function="putative packaging signals"
	/rpt_type=tandem
	3573. .5534
CDS	/note="putative semaphorin homolog; similar to Genbank Accession Number U26081"
	/codon_start=1
	/protein_id="AAC54475.1"
	/db_xref="PID:g1000717"
	/db_xref="GI:1000717"
	/translation="MAYLNATVSKPVLSISLKKYLKFGHGGEGGGLITIEFVH PAAAGTLCVSTRILMLISATTAASRPTDKRLVNLITDFGQRHPPDEPHTVLFH SLNSDVGVGNNITYLDFHSSMASTALINTSHNTRLSSTCNFTYLLANOD GLLACRSQKPSGCMILNNLTOTFLGKILAEFPSSGMLVLEFDNDITSTNLKS LSSHKFNRIAGVLEYLTDTPAMRPOQVATVHKESDDKITYEFCQNSHDFKO EPHIVPGOVGCSDDGESSLSYKKTTEFLKRLACVDYDTGRIYELDITLWAP ENMEELITGLFLSPWNPFSACVCTFYKTDIDVFKTSKENVHKKLPTPRGCKMH QHPTEFFVADRIPEVADVYQKNNAPETIOSKITYITLLVRYEYGVENATIFY LTIKIGTILHYAREDSNSTALNILEINFQKPAQIQLNDNTNKLTVSEWES EVPDLCSYVGNDCFCFMSRDPCLCTVYNNOCSEKQSVETGDPANRKLSEMGDHY APRVKHQVSIPLLSNGLSCPAVSNADYFWTKDGETEKRCVHKTKHDCILLIANS TFAENGTHCNMKEDSYTKLLEYNVTLM"
	6414. 10523
CDS	/note="putative membrane antigen; similar to p140 (PIR Accession Number S29606), and p160 (Genbank Accession Number Z26584)"
	/codon_start=1
	/protein_id="AAC54476.1"
	/db_xref="PID:g1000718"
	/db_xref="GI:1000718"
	/translation="MIRLEQLALYAPLSESALETAALRSIPGLONTVITODAYL VFEFTSRPRREHOLKLEITLFLKALRTDEQOYLIPLSALSSSTFEYGDINL PTSQSNELALLOTDRDAKEVFNVERIORLLEFSGSEOLNTHVALREILGCD LTILOPRNLFLNSSLDEYATVYIPRELKQVRHYVHGDIINLITGAASVYNNASPT AYVDSLEIRPIPTHYAFIDRMETRAVVPVGRHSGCYFRQHGCAAYVCSLNST GAGCGNGMITSFTISQPLTGQGLITEPCGWGQDLSTALITOMLKTIANVEAFK TLVAGVPLVQGVETLSPLTLREKRAMVFTSLSTGTPPLTAPTAETLDAEMERR QLLVEYGEPTYDGKGLHNPVININSESGHAHLHIALRLQMLITPFGVYHICGM QYNEATVMEFKIARACGDLGSMRSRSPARTRLKKLKNRNLNNRIITKSNLKD SAAALVIADNEVDYDSKMDIALSGCGPHILNTGPSNITVSDKNOGEYDI OYKMHMKPQSEGEDTPLAPDMSNIOLFRNLVDETDLLOVLRHPVYGCRAHIVHY DRCGNGHIAQPGVGPDPIDPLCFDSVYVNLVGDIREGSESVPRVAAMWRAARLIL PYSIISOSTRPNVLAHLPRCGPDLALALNLSGSDPTVYPGKRYHNREDAI PGSATTADPFIAGIYGVSSNIDSVDGSIROMLVNVRQIINVNRNLKAGSFTLAI GACRLIFATKAIAGFDAGSOOTPELLPNSRRESWMLFKIPEDIKAAVFDRLGC VLQPMWGQTHLGSNNITFEGLDETQOYAAENFGLVSCGARPELVNPTAEHY AGLCSDEGRHLALLPCLAFNTYQMOHONVTPGQGLPVSCKMFLRYLYNSKFEHQ HYSRLTINLRHTFNEFTQOLDPRPHARCAVNPDKNQTIPKAMDPQO"

CDS		11341..>11874	/codon_start=1 /product="major ssDNA-binding protein" /protein_id="AAC54477.1" /db_xref="pid:g1000719" /db_xref="gi:1000719" /translation="MALKHOLNHNVEDNLGSKAPIGCGFIYIPENHFNKEISILGDKYEAKAFSLPIHLGYVEAEFYPNNKAYKKIKDITTVYSKRSTFENRAIIFHNVERPESIEPSGLGSLOCKEACDLFGVVPFLDGSDVDICPVMQEDALMAVLTLEGFKERLHHGKLIIYLSQNH"
BASE COUNT	3437 a	2699 c	2362 g 3376 t
ORIGIN			
Query Match	6.5%;	Score 131;	DB 37; Length 11874;
Best Local Similarity	59.5%;	Pred. No. 7.14e-74;	
Matches 471; Conservative	0;	Mismatches 316; Indels 4; Gaps 4;	
Db	4013 CTGGCAAAACTTTTAATCTGCCTTCATTAACCAGACATGGGCGTGAAGCTTGATG	TAC	4072
Oy	375 CTGGAGAAGCTACATCATCTCTCTGGAGAGCGGGGTAGGGGCTGGCCCTGGGAC	A	434
Db	4073 TAACTCAGAGAAGACCACCTCTGGCTGTATAACAACCTAACAATCAATT	TTGGGGCC	4132
Oy	435 CAAGCCCCGGCACCCCACTCTGTGAACCTGGTAGATGGCACTGGGCCACTGGGA		494
Db	4133 AAAACTAGGCTTAAGCCCCCTTCTCACCATCATCTGGCAATGTGGTGTGGT	GACCAAA	4192
Oy	495 GATGAGAGGCTAAGCCCCCTTCAGCCCAGAGACAACTCCCTGGTTCTGTTGAAGGGGA		554
Db	4193 TGACACCTATTCACCATTAACCTCTTAC -AAGAGCCTAGTGGCTGCACAAAGTTA	GGA	4251
Oy	555 CGAGGTGTATTCACCATCTCGGAAGACAGAAATPAATGGGAAGATCCCTCG -GTTC	CCCC	613
Db	4252 GGATGGCTGGCCAAGTAGAATATACAGAGATGACCCGCAATGGACCGGCACAGTTG		4311
Oy	614 GCATCCGGGGGAGAGTGAAGTGAAGTACACATGATACGTATGAGAAACCAAGTTCA		673
Db	4312 TCACAGCAACAGCTGTGCATAAAAATGAATCTTATGATGATAAAATCTACT	TTTTCTTC	4371
Oy	674 TCAAAGCCACCATGTGTGACCAAGACCGGCTTACAGTAGAAGATCTACTACT	CTTCC	733
Db	4372 AAGAAAGACGCACAGTACTTCAACAGTTTCCACATACGTACTAGATGGGGCAGG		4431
Oy	734 GAGAGGAATATCTGACAAGAATCCTGAAGGCTCTCTCAATGATGTCCGTGTGGCCAGT		793
Db	4432 TGTGCTAGTATGATCAAGGTGGGAGAGCTCCTGTGTCTACAAAGTGGAACCTTT		4491
Oy	794 TGTGAGGGGGGACCAAGGTGGGAAAGTTACATCTGATCTCCAAAGTGAACATTTTC		853
Db	4492 TAAAGCCAGAGCTGGCGTGTAGACTATGATCTGGAAGATCTCAATGAGCTACAAG		4551
Oy	854 TGAAGCCATGTGTATGACAGTATGCTGTGCACCAACAAACTTCAACAGGCTGCAAG		913
Db	4552 ATATTTCAT -CTGGCAAGCCCCAGAGAAGAGCTGGGAAGAGACTCATATATAGATT		4610
Oy	914 AGCTTCTCTCTCTCCCTGACCCCAAGCCGACAG -TGAAGGACACCAAGGCTATGTGTT		972
Db	4611 TTTTGAAGCCCGTGAATTTTCTGGCGGTGTGTGTTTACTGTAAAGACATTGACAT		4670
Oy	973 TTCTCCAAACCCCTGGAATCTACTCAAGCCTCTGTGTGTTTCTCCCTGGGATATGACAAG		1032
Db	4671 GTGTTTAAAGACATCAAGTTAAAAAATATATCATATAAACTCCCAACCTAGACCAAGG		4730
Oy	1033 GTCTTCGCTACTCTCTCACTCAAGGGCTAACCACTCAACCTTCCCAACCCGCGGCTGGC		1092
Db	4731 CAAATGCAAGAAACCAATACATATGTTCCCAAGAAACCTTCAAGTGTGGACAGATAT		4790
Oy	1093 AAGTCCCTCCAGACCAACCAATATCCCAAGAGACCTTCCCAAGGTGGCTGACCGTCAAC		1152
Db	4791 CCAGAAGTTGC 4801		
Oy	1153 CCAAGAGGTGC 1163		

RESULT 8 AF005370 130608 bp DNA VRL 21-AUG-1997
LOCUS Alcelaphine herpesvirus 1 L-DNA, complete sequence.
DEFINITION AF005370
ACCESSION AF005370
MID G2337967
VERSION AF005370.1 GI:2337967
KEYWORDS
SOURCE .
ORGANISM wildebeest herpesvirus.
Alcelaphine herpesvirus 1
viruses: dsDNA viruses, no RNA stage; Herpesviridae;
Gammaherpesvirinae.
REFERENCE
AUTHORS Ensser, A., Pflanz, R. and Fleckenstein, B.
TITLE Primary structure of the alcelaphine herpesvirus 1 genome
J. Virol. 71 (9), 6517-6525 (1997)
MEDLINE 97404659
REFERENCE 2 (bases 1 to 130608)
AUTHORS Ensser, A., Pflanz, R. and Fleckenstein, B.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-1997) Institut fuer Klinische und Molekulare
Virologie, Friedrich-Alexander-Universitaet Erlangen-Nuernberg,
Schlossgarten 4, Erlangen 91054, Germany
FEATURES
source
1. 130608
/organism="Alcelaphine herpesvirus 1"
/strain="C500"
/specific_host="Connochaetes taurinus taurinus"
/db_xref="taxon:35252"
/lab_host="MDK bovine epithelial kidney cells; ATCC
CCL-22"
918. 1211
/note="putative ORF"
/codon_start=1
/evidence=not_experimental
/product="A1"
/protein_id="AAC58052.1"
/db_xref="PID:g2337968"
/db_xref="GI:2337968"
/translation="MRLHFLYRSIEFMYYKRSRHKIFCITLSIVELTAHKKRPARNP
IDFBSRIGPCFVKRSRQHLIVNOLMSPCRNCKSLTSAOCCITANEDTTK"
1248. 1254
/evidence=not_experimental
complement(1300..1305)
/note="spliced cDNA detected by 5'RACE-PCR"
/codon_start=1
/evidence=experimental
/product="A2"
/protein_id="AAC58053.1"
/db_xref="PID:g2337969"
/db_xref="GI:2337969"
/translation="MSQNSSENPSRRKRYVKMCDITEOKERRRSINRASKRNF
RRRIPEOQEKLINIKYENSRKRCQVEKRKDEIRLRWLVHKKCTLLQNTGPPE
PVKVENSLKMOCAFAFLNLDQQTNNINLETYSGNNTINGFAAATATHTNCEYK
TLANNTNFEALNCEVLPSTLALDILSIDMNNLYN."
3492. 5453
/note="A3: AHV-sema, similar to Vaccinia A39"
/codon_start=1
/evidence=not_experimental
/product="semaphorin homolog"
/protein_id="AAC58054.1"
/db_xref="PID:g2337970"
/db_xref="GI:2337970"
/translation="MAYLNATYSKPVISLSISKVYLFENCGGSGCIGLITEVIVH
PAMGGLVYSIRLWLSITAKSRFIDKPRILVNLTDGFGQARFEPQEPHYVLFH
SLNSDDVYGGNNITLYLDFEHAHSSNASTALINTSTHNRHLSSTCEMTILNHTQD
GLACGTSQKSCWLINNLITQFLGPKLAPFSPSSGNLYLFPQNDYISTINLYKS
LSGSHKFRRIAGOVELYTSDTAMHRPOFGAARVAKNESYDKIYEFPOENSHSPFO
EPHTYPRVGQVSSDQGGSSLSYKMTTELKARLACVDYDGRIRYNLEQDFTIQAP
ENSMEETLTYGFLSPMNSAYCAVTVKADIDHVFRTSKTKLHRLPFRPEQCKNH
QHPETETQVADRYPEVADPYQKNNAPFIQSKYITKLLVYVEVGGVMAQITFY
LTIGTITHIYVEDSNSTALNITLLEINPFOKPAPIQNIILDNLTNLKLYNSEWEVS

polyA_signal
CDS
EYPLDLCVYGNDCFCSEFMSRDLCTWYNNTCSFKORYSEVTGPGANRTLSEMGDHY
APTVYKHQVSIPLNSYSLSCPAVSNHADYFPTKGFTEKRQHVATHKNDICLLANS
TQATNGTHYCNKNEKDSVYKLLVNTLM"
5590. 5597
5732. 6097
/note="ORF containing putative signal peptide: detected by
5'RACE-PCR"
/codon_start=1
/evidence=experimental
/product="A4"
/protein_id="AAC58055.1"
/db_xref="PID:g2337971"
/db_xref="GI:2337971"
/translation="WVAOLYHHLITLISIVYIFPVNAPRLTEDEDEDNSTEVDF
LDSDFQPPRRSSCNCRTREERITLLELRKRALQIINGINMETDEUPEYTPPTMTT
PLRTPLDTSPPVLPSPALP"
6123. 6128
/evidence=not_experimental
6333. 10442
/note="ORF03; similar to H. salmieri and EBV2 ORF3, similar
to ORF75"
/codon_start=1
/evidence=not_experimental
/product="tegument protein"
/protein_id="AAC58056.1"
/db_xref="PID:g2337972"
/db_xref="GI:2337972"
/translation="WITREQLLAYTAPSALETAIALNSIPGLQNVIIOTDAYL
VTFSPRPRLEHQKLETLFIKALTEPBOOYLPIPLASLSSFTFVGPDIRL
PTSQNEELTALIQTRDAREKAYENYRICHRLCFSSGQOLNHTHALREILCGD
LILYQFPRNELENSSLDIFYATYPPRFKDYVRHVGDINDLITGAASNYNMSPT
AVTDEIRIPRIPTYHVFRIDRMEIRIDRVGFGHSGRGROROGGAAYMSTNST
GATGCLQKGMTSPISOPTLQGLITTPKCKWGQOESATRIQMLTRYANTVEFAK
TLNTAGVPLVGSEVTLSPTLREKRMAFTGSLSTGLPPYILKPEPTATLDAEMNR
OLLIVEGYPDYGGKGLHPVNNINSESGRAHLIQLRLQMLITTPRKYLITCCDW
QYNEKVTMFKIAKAGDGLSMSARSLPAHRLRLKKNFNLYNRI INKSLKVD
SAALVVIADNEYODVSKMDIALSGMCCPHILIGNLPNSNTIYISDKNOYGEIVDI
OYKMHMPKQPSGTEPDPPLADMSNIOLEFKMLDTEDLLOVLRHPYGCXAHVHV
DRCGNGHIAQOPGVCPDPLCDPSVYTHNVDDDIRGMSYRVMAADRVARRLI
EYQTSPELDITLANNLGLTHIRPEHGVYTESKRYGMCIGTEGTTTQDDPL
GITLIVSECTNCLIGPEVNEEELIGTISVPEIYHRAHENSIMAAKDFCSMNF
GFVNASANGCILLRSVLANPACVVPSPSLKPYFKPSAISILRVNLHTFHSGLI
CCMASGASGETFPTPSOLNRLQFMLIVAENMLALSHGSDGLICAVCEMFFAG
GASARLIIHDEDEBPPLFSETPEGFVYVNAIDVAAIIARANLYNEVCIOIGEVVS
DTEVYHONTOLLISVPSRIKHNHTFESKSDLVLYKEDQVLPETSQNGEYHLYTD
PYSTISQSTTPNVLVHLIPCGGYPDALALITNSGSPDVTVPCKRYRNNREDAA
PGSPPADDFIAGIVLYGSSNIDSDVGDSTIRQWLVNVRQVINDVARNLKAGSTFLAI
GOLACRILIFATKAIGFDAGSOOTPELLPNASRRYESRLNFKIPEDTKAFAFRLRC
VLPCWVGTHLGFSSHNNTTFEGDLETRQOVATNGPLVQSGPAREVPLNPTAEHPY
AGLCEGDEGRHLALLEPDLAFNTWQWONOTGPGQELPVSPKRLMFFRLNWSKFIQ
HYRSLRNLNHTFNEQQLDFHADRGAVPHDNPQYITMDAMDSQ"
11134. 11139
/evidence=not_experimental
11260. 14643
/note="ORF06"
/codon_start=1
/evidence=not_experimental
/product="major ss DNA binding protein"
/protein_id="AAC58057.1"
/db_xref="PID:g2337973"
/db_xref="GI:2337973"
/translation="MALKHOLNHAVEDNLGSKAPIGCGFIYIPETHFNKETSILG
DKYAEAGFSLPLHGTVEAPVPNNKAYKIKDNTTYSKLTFTYIRAIIFHNVEK
PESTFSGDGLCKEKQDMLGIVPFPPLGGSDVDSDIPQWQEDKSTFTYIRAIIFHNVEK
FERKRLHGLKLLYLSQMSHVMINRTEYVRLELYXEDLFTKSSLRRLYLPAAVSELYL
TLYLSLQSLQSVNVAASIVEAIIQOQFHXDKMKAKLVSFYKPLATYAGKQALTIMVID
AAVELEIGSYLSIFEPAQOEKTYQODYSMOIFASCETDSRLPALKMAKLAQIITHI
AOLSTNSIYVNVNVAOAPLPNSKVRNVAVNSYVLOHGLANLEETLFEDEGSAFPG
APASIDSGSTTOLHAYAAFSNLAEMKYIYQFOGHOKSTINPARNITEYSGAA
NSPVCISLCSGCCPVCINTLYLRIKDRFPPLVQGSRRDPYITITIVFENLDFLGNF
ASPRDKQEDQWQETEPRTYTYLQNLQTELEAAGLVDSVADEAGAGGSGSNMLEKF
VRTSDSDLSLDAANAFTINIMINNVNFKESIGSVHVIQYONNTYQWQAFCSLMLL
YRSILITIQDIALPISTVSESNPAGQYKREMLKLYHQGLMTNFKSFIDKGVITG

4052 AAACCTAGGCTTAGCCCCCTTCTCACCATCATCTGGCAATCTGGTCTGTTGACCAGAA 4111

1086 yyyyyyyyyyyyyyyyyy 1146

[illegible]

QY 1661 TTTCCTGGCCCAACTCTGCTACTGAGCTGCCCATGAAATCCCGCAGCCCA 1720
Db 1146 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1205
QY 1721 CTTACTATGCGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1780
Db 1206 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1265
QY 1781 CCAACTGACATCTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1840
Db 1266 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1325
QY 1841 AGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900
Db 1326 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1385
QY 1901 GCATCATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1960
Db 1386 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1435
QY 1961 GGGTGTGCGCCACACATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2010

RESULT 10
LOCUS AR024229 965 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 22 from patent US 5795961.
ACCESSION AR024229
NID 93977523
VERSION AR024229.1 GI:3977523
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 965)
AUTHORS Wallace,T.Paul, Harris,W.J., Carr,F.J., Old,L.J., Welt,S. and Kitamura,K.
TITLE Recombinant human anti-Lewis b antibodies
JOURNAL Patent: US 5795961-A 22 18-AUG-1998;
FEATURES
source 1. 965
BASE COUNT 192 a 170 c 226 g 205 t 172 others
ORIGIN

Query Match 1.9%; Score 38; DB 25; Length 965;
Best Local Similarity 13.8%; Pred. No. 7.67e-07;
Matches 20; Conservative 71; Mismatches 53; Indels 1; Gaps 1;

Db 802 DYTTSYMGVNRGRGMDYGGYTNNGRGRVYTMADTSSNSRSTYATADTAVYCYGRS 861
CP 1216 ATTAGAGTGAACAATGCGCTCTCAGAGGCCCATGGGCTCCACCTCTCGCCACCT 1157

Db 862 YOSDGDYNGGTTVYSSHTVADMTSSSSASVDRTRTSSSTTHGNGTYYWYKAKAY 921
CP 1156 CTGGGTGAGCGTCAGCGCCTCTGTAAGGCTCTGTGGGTATCGCTGCTGAGAGCG 1097

Db 922 RVSNRSGVSRSGSGCTDYTTSSDA 946
CP 1096 ACTTGCCAGGC-CGCGGCTTGGGAA 1073

RESULT 11
LOCUS MVU87256 1056 bp DNA MAM 02-JAN-1999
DEFINITION Mustela vison GT dinucleotide repeat, chromosome 1q.
ACCESSION U87256
NID 94099442
VERSION U87256.1 GI:4099442
KEYWORDS
SOURCE American mink.
ORGANISM Mustela vison
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Carnivora; Fissipedia; Mustelidae; Mustela.
REFERENCE 1 (bases 1 to 1056)

AUTHORS Brusgaard,K., Shukri,N.M., Malchenko,S., Koroleva,I. and Lohi,O.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1997) Breeding and Genetics, Danish Institute of Animal Science, Blichersalle 825, Tjele 8830, DK
FEATURES
source 1. 1056
/organism="Mustela vison"
/db_xref="taxon:9667"
/chromosome="1"
/map="1q"
/note="primers: 1167F: agccctgcatactactctt, 1167R: gagatctaccgctgttag"
98 119
/standard_name="1167F"
primer_bind complement(300..320)
primer_bind /standard_name="1167R"
BASE COUNT 211 a 221 c 210 g 225 t 189 others
ORIGIN

Query Match 1.9%; Score 38; DB 23; Length 1056;
Best Local Similarity 19.0%; Pred. No. 7.67e-07;
Matches 23; Conservative 59; Mismatches 38; Indels 1; Gaps 1;

Db 559 GMSCKMTWTGSCCWSCHRDVBMKSKWCKKSCSTGKDKMSGCAVCGRRRCR 618
CP 128 GAGGCGGCGCGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 69

Db 619 YMMRMVCGSGTWARCCCDKRSKSGDNKH-CKSRKRYKMDRYHBCKSMCAVKGWSA 677
CP 68 GCCGCGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9

Db 678 G 678
CP 8 G 8

RESULT 12
LOCUS MUSIGF2R 2582 bp DNA ROD 15-MAR-1999
DEFINITION Mus musculus Insulin-like growth factor 2 receptor (Igf2r) gene, exon 1 and partial cds.
ACCESSION L06445
NID 9293379
VERSION L06445.1 GI:293379
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2582)
AUTHORS Barlow,D.P., Kubicka,P., Liu,C.G., Kafri,T., Razin,A., Cedar,H. and Stoger,R.
TITLE Maternal-specific methylation of the imprinted mouse Igf2r locus identifies the expressed locus as carrying the imprinting signal
JOURNAL Cell 73 (1), 61-71 (1993)
MEDLINE 93214996
FEATURES
source location/Qualifiers
1..2582
/organism="Mus musculus"
/strain="C57/BL6"
/db_xref="taxon:10090"
/chromosome="17"
<1215..>2582
/gene="Igf2r"
1215..>1348
/gene="Igf2r"
/codon_start=1
/product="insulin-like growth factor 2 receptor"
/protein_id="AAA37921.1"
/db_xref="PID:g553942"
/db_xref="GI:553942"
/translation="MRAVQLGVPVSGPRVALLPPLLILLAAAGSAQAQAVLDALC"
<1215..1348
exon

EDDTSITPTDFYINICOPLPNMGVPCPAGASVCKPVYDGPIDIGRTGPIENPVA
NEVYLFESTPTCLADKXMYNTSLIAFHCRGISMGTPKLTIRNDDCPREMEPIYV
PDEVYTOGCAVTDOLLSFNLJSTSTSTFTVTRAHAYTISGVCCTAADIQEGCKDG
GCVLLSGSKGASFGRLASMOQDRIHRQDEAVILSTYNGDPCPETEDGEPCVFPYKRG
ETVDECVLEGRKLMCKSTANIYDRHDEWFCRPSNSHRMSA1IFTCDENEDIGREVE
SEDRGCEVTEFEMKTRVYCPKMECKEYOKHKTVDRLSLSTGSMDFVHGSNSYFN
LCORVYKGPLDCSERASICKSATGOVVLGIVHTOKLEIYDIDETVLYTSKRGSGGN
KTASVIELTCAKTYGAPAFKRPDIDSCITYEYWSRAACAVRPOEVNMYGTLNPNY
TKSFSLGEIYFKLFSASGDMRSGNDNYLFIQLSISISSNPACSGANICQVKNPD
HESRKGVSMDTKYVQDGLDYFTSSSTGCKDKTSVSTIFFHCDPLKDGPER
SHETRDQYLFWSITAVCPGLVDFDDENAVPEYKGLSERQAVGAVLSLVALTGC
LLALLHKKEREYVINKLTNCCRRSSGVSYSKVSKEETDENETEMLEEIQVPA
PRLKDGQENGHITTKYKAEALTSLHGDDQSEDEVLTIPVAVHTGREGAVESSOP
LRNFORKLVKEREGRMGLVGEKARGRKFRPGQRKPTTPAKLVSFHDDSDIDLHT"

BASE COUNT 2224 a 2125 c 2369 g 2092 t

ORIGIN

Query Match 1.9%; Score 38; DB 32; Length 8810;
Best Local Similarity 83.9%; Pred. No. 7.67e-07;
Matches 47; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 135 CTGCTGCCGCTCTCTGCTGCTGCTGCTGCGCGCGGCTCCGCCAGGGCCA 190
OY 85 CTGCGGCTGCGGCTCTGCTGCTGCTGCGCGCGGCTCCGCCAGGGCCA 140

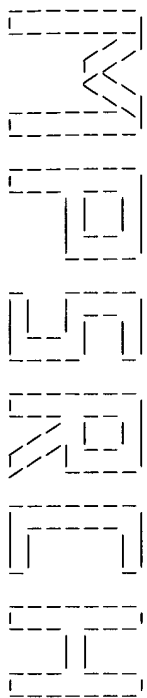
RESULT 15
LOCUS 128278 215 bp DNA PAT 30-OCT-1996
DEFINITION Sequence 5 from patent US 5569830.
ACCESSION 128278
NID g1819054
VERSION 128278.1 GI:1819054
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 215)
TITLE Bennett, A., Labavitch, J.M., Powell, A. and Storz, H.
JOURNAL Plant inhibitors of fungal polygalacturonases and their use to
FEATURES control fungal disease
PATENT: US 5569830-A 5 29-OCT-1996;
LOCATION/Qualifiers
source 1..215

BASE COUNT 15 a 8 c 25 g 26 t 141 others
ORIGIN /organism="unknown"

Query Match 1.8%; Score 37; DB 25; Length 215;
Best Local Similarity 13.9%; Pred. No. 3.06e-06;
Matches 29; Conservative 84; Mismatches 94; Indels 2; Gaps 2;

Db 6 MSSSVYSRTASCNDKAKKDGNTSSWTTDCNRTWGVCDTDTTYRVNNDGHNKYSAN 65
OY 244 CCAGGCAGCTCTCTGTGTGGAGAGAGCTGCAAGGTCTTGTACTTGTCTCC 303
Db 66 YNYGNVGAATHTYHTVNSGADSKTVTDYSYNASGTSSNGTDCNRSAGADSYSSKT 125
OY 304 GAGGCAAGAACGATCTGTGCGACGCTGAATCTCGCTCCACAAAGGGTCTGTGTG 363
Db 126 AMTSRNFTGTANNVADSRNMGDASVSGDKNTKKHAKNSADGKYVSKNNGDRNNRYGTGT 185
OY 364 GAT-AAAGCGGAGCTGACGAGACTACATCTCTCTGAGAGGCGGAGTGAGGGGC-TGC 421
Db 186 KSNVSNKCGGKRDVSSYANNKCCGSSC 214
OY 422 TGGCTGTGTGACCAACGCCGCGACCCC 450

Search completed: Tue Aug 3 22:24:47 1999
Job time : 4032 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

Mpsrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Aug 3 23:27:13 1999; Maspar time 443.07 Seconds
972.040 Million cell updates/sec

Tabular output not generated.

Title: >US-09-240-410-1
Description: (1-2010) from US09240410.seq

Perfect Score: 2010
N.A. Sequence: 1 ATGACGGCTCTCCGCGCCG.....TGTCACATAGAGGCTCCG 2010
Comp: TACTGCGAGGAGGCGGCGC.....ACGAGTGATCCGAGAGGGC

Scoring table: TABLE default
Gap 6

Mismatch STD : Dbase 0; Query 0

Searched: 271905 seqs, 107135622 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

n-geneseg35
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40 41:part41 42:part42 43:part43
44:part44 45:part45 46:part46 47:part47 48:part48
49:part49 50:part50 51:part51 52:part52 53:part53
54:part54 55:part55 56:part56 57:part57 58:part58
59:part59 60:part60

Statistics: Mean 9.556; Variance 6.017; scale 1.588

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description	Pred. No.
1	47	2.3	190 32	T76452	Chymase antisense oli	6.53e-12
2	44	2.2	178 32	T76405	Human endothelin-1 an	3.05e-10
3	44	2.2	204 1	N81164	Base substituted E.co	3.83e-09
4	42	2.1	91 9	O51746	Oligonucleotide probe	1.08e-09
5	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
6	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
7	38	1.9	114 12	O70467	Generic DNA sequence	5.46e-07
8	39	1.9	114 12	O70469	Generic DNA sequence	1.60e-07
9	38	1.9	114 12	O70465	Generic DNA sequence	5.46e-07

ALIGNMENTS

RESULT	ID	Score	% Match	Query Length	DB ID	Description	Pred. No.
1	T76452 standard; DNA; 190 BP.	47	2.3	190 32	T76452	Chymase antisense oli	6.53e-12
2	T76452; 16-SEP-1997 (first entry)	44	2.2	178 32	T76405	Human endothelin-1 an	3.05e-10
3	Chymase antisense oligonucleotide.	44	2.2	204 1	N81164	Base substituted E.co	3.83e-09
4	Asthma; airway epithelium; adenoviral free; cystic fibrosis; chronic obstructive pulmonary disease; bronchitis; ss.	42	2.1	91 9	O51746	Oligonucleotide probe	1.08e-09
5	Synthetic.	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
6	MO9640162-A1.	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
7	PD 19-DEC-1996.	38	1.9	114 12	O70467	Generic DNA sequence	5.46e-07
8	06-JUN-1995.	39	1.9	114 12	O70469	Generic DNA sequence	1.60e-07
9	07-JUN-1995.	38	1.9	114 12	O70465	Generic DNA sequence	5.46e-07
10	PI Metzger WJ, Nye JW; (UYEC-) UNIV EAST CAROLINA.	47	2.3	190 32	T76452	Chymase antisense oli	6.53e-12
11	PI Metzger WJ, Nye JW; (UYEC-) UNIV EAST CAROLINA.	44	2.2	178 32	T76405	Human endothelin-1 an	3.05e-10
12	DR 97-051871/05.	44	2.2	204 1	N81164	Base substituted E.co	3.83e-09
13	PT Treatment of airway diseases such as asthma - by topically applying adenoviral free antisense oligo:nucleotide to airway epithelium of	42	2.1	91 9	O51746	Oligonucleotide probe	1.08e-09
14	PT Subject	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
15	PS Example 5; Page 40; 71pp; English.	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
16	CC A method for treating airway disease in a subject has been produced, CC which involves the topical administration of an essentially adenoviral free antisense oligonucleotide (ON) to the airway epithelium of the	42	2.1	91 9	O51746	Oligonucleotide probe	1.08e-09
17	CC subject. The present sequence is an antisense oligonucleotide specific CC for chymase, targeted at the initiation codon. The method can be CC used to treat airway diseases such as cystic fibrosis, asthma, chronic	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
18	CC obstructive pulmonary disease, bronchitis and other airway diseases CC characterised by an inflammatory response. By eliminating adenoviral	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
19	CC from the antisense ON, its liberation upon antisense degradation is CC prevented, thereby preventing adenoviral-induced bronchoconstriction	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
20	CC In patients with hyper-reactive airways.	43	2.1	204 1	N81164	Base substituted E.co	1.08e-09
21	Sequence 190 BP; 1 A; 41 C; 58 G; 36 T;	38	1.9	114 12	O70467	Generic DNA sequence	5.46e-07
22	Query Match	39	1.9	114 12	O70469	Generic DNA sequence	1.60e-07
23	2.3%; Score 47; DB 3; Length 190;	38	1.9	114 12	O70465	Generic DNA sequence	5.46e-07


```
Matches 5: Conservative 34; Mismatches 73; Indels 0; Gaps 0;
Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
QY 1392 CATCAAGACCATGCTGCTGAGTGTGAGCGAGAGACGTATGTGCTCCAGTGGGA 1451
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
QY 1452 GGTGAGCGAGGTGCTGAGCCTGTGTGAGGTATGCGGGGGCGCCAC 1503

RESULT 8
ID 070469 standard; DNA; 114 BP.
AC 070469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KM TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc-feature 55..60
FT /tag= a /note= "this sequence represents '2'; 2 can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PI WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
PI WPI: 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
CC 070469 is a generic DNA sequence used to generate random TSAR peptide
CC This generic formula can be represented as follows: X(TGC)(NNB)10-
CC (TGC)(NNB)62(NNB)(TGC)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in 070465-68. Other specific peptides
CC generated by these generic sequences are shown in 865150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active. They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or compns. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP: 0 A; 4 C; 4 G; 4 T;

Query Match 1.9%; Score 39; DB 12; Length 114;
Best Local Similarity 8.8%; Pred. No. 1,60e-07;
Matches 10; Conservative 31; Mismatches 72; Indels 0; Gaps 0;
Db 1 tgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 60
CP 1883 TCCCACTGCTGAGCCCTGCGGAGATGAGACCTCTGCGCCCTCGCAGAGATGCGCG 1824
Db 61 nbnbnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 113
```

```
CP 1823 TACTGCTGCCCCCTGAGCTTCTGATGACAGATGATGAGTGGGCTGTGCTG 1771

RESULT 9
ID 070465 standard; DNA; 114 BP.
AC 070465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KM TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc-feature 55..60
FT /tag= a /note= "this sequence represents '2'; 2 can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PI WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
PI WPI: 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
CC 070465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)112(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in 070466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC 865151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP: 0 A; 2 C; 2 G; 2 T;

Query Match 1.9%; Score 38; DB 12; Length 114;
Best Local Similarity 5.4%; Pred. No. 5,46e-07;
Matches 6; Conservative 33; Mismatches 73; Indels 0; Gaps 0;
Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
CP 1266 CTCGCCGTGCTGCTGCTGATGCGGTGAGCGGCCACTTCTGAGTGTGATTTAGAGTG 1207
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
CP 1206 GACCAATGGCTCTTCACAGAGCGCCCATGCGCTCCACCTCTGCGCCACCTCT 1155

RESULT 10
ID 070467 standard; DNA; 114 BP.
AC 070467;
```


Query Match

Best Local Similarity 46.4%; Pred. No. 6.18e-05;
Matches 26; Conservative 20; Mismatches 10; Indels 0; Gaps 0;

Dn Db 21 ggccttcgagcggcggtccgcbgagcgtcgcbbgacbbcbtgcgbgagbgcgtctcbb 76
|||||:|:| :|||:| |:: |||: ::| |:: ||:|:| | |:
YQ 77 GGCTTCGGCGTCCGCGGTCTGCTGCTGCTGCTGCTGCGGCGCGCGCTCCGCC 132

RESULT 14

ID Q70469 standard; DNA: 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR: locally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60 /tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"

PN WO9418318-A.
PD 18-AUG-1994.
PE 01-FEB-1994;
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.

Query Match

Best Local Similarity 46.4%; Pred. No. 6.18e-05;
Matches 26; Conservative 20; Mismatches 10; Indels 0; Gaps 0;

Dn Db 21 ggccttcgagcggcggtccgcbgagcgtcgcbbgacbbcbtgcgbgagbgcgtctcbb 76
|||||:|:| :|||:| |:: |||: ::| |:: ||:~|:| | |:
YQ 77 GGCTTCGGCGTCCGCGGTCTGCTGCTGCTGCTGCTGCGGCGCGCGCTCCGCC 132

RESULT 13

ID T76219 standard; DNA: 89 BP.
AC T76219;
DT 12-SEP-1997 (first entry)
DE Human IL5 antisense oligonucleotide.
KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; interleukin; ss.
OS Synthetic.
PN WO9640162-A1.
PD 19-DEC-1996.
PE 06-JUN-1996; U093306.
PR 07-JUN-1995; US-474497.
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger WJ, Nyce JW;
DR WPI: 97-051871/05.
PT Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS Claim 5: Page 30: 71pp; English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for the human IL5, targeted at the initiation codon. The method can be
CC used to treat airway diseases such as cystic fibrosis, asthma, chronic
CC obstructive pulmonary disease, bronchitis and other airway diseases
CC characterised by an inflammatory response. By eliminating adenosine from
CC the antisense ON, its liberation upon antisense degradation is
CC prevented, thereby preventing adenosine-induced bronchoconstriction in
CC patients with hyper-reactive airways.
SQ Sequence 89 BP; 0 A; 20 C; 25 G; 18 T;

PA	(YUNC-) UNIV NORTH CAROLINA.
PI	Fowlkes DM, Kay BK;
PT	WPI: 94-279739/34.
PR	Identifying proteins or peptide(s) which bind a ligand - by
PS	screening a recombinant vector library expressing fusion proteins
PP	comprising a binding domain and an effector domain
PB	Disclosure; Page 35; 255pp; English.
CC	Q70465 is a generic DNA sequence used to generate random TSAR peptide
CC	This generic formula can be represented as follows: X(TGC)(NNB)10-
CC	(TGC)(NNB)6(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC	sites (X is not the same as Y) that are not specified further. This
CC	sequence generates peptides that are cloverleaf in structure. Other
CC	generic sequences are shown in Q70465-68. Other specific peptides
CC	generated by these generic sequences are shown in R65150-94. TSARs are
CC	concatenated heterofunctional proteins or peptides, comprising at least
CC	two functional regions - a binding domain with affinity for a ligand and
CC	a second effector peptide portion that is chemically or biologically
CC	active.They may further comprise a linker peptide between the 2 domains.
CC	The oligonucleotides are also designed so that the expressed peptide
CC	contains 2 or 4 cysteine residues positioned in, or flanking, the
CC	unpredicted or variant residues. These residues confer some degree of
CC	conformational rigidity to the peptides. The TSARs or compns. comprising
CC	a TSAR binding domain can be used in vivo to deliver a chemically or
CC	biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC	or enzyme, to the specific target or on the cell. They can also replace
CC	the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC	and therefore circumvent the need for complex methods of hybridoma
CC	formation or in vivo antibody production. The TSARs are easily
CC	characterised and have designed activity allowing direct and rapid
CC	detection in a screening process.
SQ	Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;
Query Match	1.8%; Score 37; DB 12; Length 114;
Best Local Similarity	8.3%; Pred. No. 1.85e-06;
Matches	9; Conservative 30; Mismatches 70; Indels 0; Gaps 0;
Db	6 bmbnbnbnbnbnbnbnbnbnbnbtgcnnbnbnbnbnbnbnbnbnbnbnbnbn 65
QY	1858 TACTTCGCGAGGCTCAGCATGGCGAGCTGTCGCCCGAGGAGCATCATGCGCGACAC 1917
Dd	66 btgcnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbtgc 114
OY	1918 CTGTGGTGATGCTGCTGTGCCCTGGCGGCTCTCTGTGGTGGGGTGC 1966
RESULT	15
ID	Q70468 standard; DNA; 114 BP.
AC	Q70468;
DT	05-APR-1995 (first entry)
DE	Generic DNA sequence to generate a random TSAR peptide library.
KM	TSAR: totally synthetic affinity reagent; synthetic; binding domain;
KW	effector domain; concatenated heterofunctional protein; linker;
KW	direct; rapid; detection; screening; treatment; genetic; ss.
OS	Synthetic.
FH	Key Location/Qualifiers
FT	misc_feature 55..60 /tag= a
FT	/note= "this sequence represents 'Z'; Z can be a
FT	sequence of 6, 9 or 12 nucleotides (see
FT	comments)"
PN	WC09418318-A.
PD	18-AUG-1994.
PE	01-FEB-1994; U00977.
PR	01-FEB-1993; US-013416.
PR	30-DEC-1993; US-176500.
PR	31-JAN-1994; US-189331.
PA	(YUNC-) UNIV NORTH CAROLINA.
PI	Fowlkes DM, Kay BK;
PT	WPI: 94-279739/34.
PR	Identifying proteins or peptide(s) which bind a ligand - by
PS	screening a recombinant vector library expressing fusion proteins
PP	comprising a binding domain and an effector domain
PB	Disclosure; Page 35; 255pp; English.
CC	Q70465 is a generic DNA sequence used to generate random TSAR peptide
CC	This generic formula can be represented as follows: X(TGC)(NNB)10-
CC	(TGC)(NNB)6(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC	sites (X is not the same as Y) that are not specified further. This
CC	sequence generates peptides that are cloverleaf in structure. Other
CC	generic sequences are shown in Q70465-68. Other specific peptides
CC	generated by these generic sequences are shown in R65150-94. TSARs are
CC	concatenated heterofunctional proteins or peptides, comprising at least
CC	two functional regions - a binding domain with affinity for a ligand and
CC	a second effector peptide portion that is chemically or biologically
CC	active.They may further comprise a linker peptide between the 2 domains.
CC	The oligonucleotides are also designed so that the expressed peptide
CC	contains 2 or 4 cysteine residues positioned in, or flanking, the
CC	unpredicted or variant residues. These residues confer some degree of
CC	conformational rigidity to the peptides. The TSARs or compns. comprising
CC	a TSAR binding domain can be used in vivo to deliver a chemically or
CC	biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC	or enzyme, to the specific target or on the cell. They can also replace
CC	the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC	and therefore circumvent the need for complex methods of hybridoma
CC	formation or in vivo antibody production. The TSARs are easily
CC	characterised and have designed activity allowing direct and rapid
CC	detection in a screening process.
SQ	Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;
Query Match	1.8%; Score 37; DB 12; Length 114;
Best Local Similarity	8.3%; Pred. No. 1.85e-06;
Matches	9; Conservative 30; Mismatches 70; Indels 0; Gaps 0;
Db	6 bmbnbnbnbnbnbnbnbnbnbnbtgcnnbnbnbnbnbnbnbnbnbnbnbnbn 65
QY	1858 TACTTCGCGAGGCTCAGCATGGCGAGCTGTCGCCCGAGGAGCATCATGCGCGACAC 1917
Dd	66 btgcnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbtgc 114
OY	1918 CTGTGGTGATGCTGCTGTGCCCTGGCGGCTCTCTGTGGTGGGGTGC 1966

This Page Blank (uspto)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 459)

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nh.gov
This clone is available royalty-free through LNLN ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 2686 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 446.

Location/Qualifiers
1. .459

unwired-scanned cDNA was size selected, ligated to Eco adapters (Pharmacia), digested with Not I and cloned in

the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and

M.Fatima Bonaldo."
/db_xref="taxon:9606"

```
/clone="IMAGE:1759553"
/clone_lib="Soares_placenta_8tc9weeks_2nbHP8tc9w"
```

```
/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
```

```

/lab_host="DH10B (ampicillin resistant)"
BASE COUNT      87 a      134 c      151 g      87 t

```

ORIGIN

Query Match	22.28;	Score 447;	DB 24;	Length 459;
Best Local Similarity	98.78;	Pred. No. 0.00e+00;		

Matches	453;	Conservative	0;	Mismatches	6;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	----	--------	----	------	----

Dd 1 GCGGGGGGGCCCTCCGGCCCAAGGGGGACCCTAAGGAGGGAGCCCCGCATCTTCCCGCTGG 60

QY 115 GCGGCGCGCCTCCGCCCAGGGCCACTAAGAGCGGACCCCGCATCTCGCCGCTGG 174

D6 61 AAGSCCATGTAGGGCAGACCGGGTGCATTTGGCCAGACTGAGCCGACACGGTCTT 120

QY 175 AAAGSCATGTAGGGGACAGACCGGGTGGACTTTGGCCAGACIAGAGCCGACACGGTGCTT 234

D5 121 TTCCAGGAGCCAGGCTCTCTGTGTGGGTGGGAGGACGTGGCAAGGCTTACCTTTT 180

QY 235 TTCCACGAGCCAGCCAGCTCTCTGTGTGGGTGGGAGGACGTGGCAAGGTCTACTCTTT 294

D5 181 GACTTCCCCGAGGCGCAGACGCATCTGTGGCGCAGGTGATATCCGCTCCACCAAGGG 240

295 GACATCCCCGAGGCGCAAGACGCATCTGTGCGCAGGTGATATCGGCTCCACCAAGGGG 354

D5 241 TCCGTCTGGAAAGCGGCACTGTGAGAACTACTATCACTCTCCCTGGAGAGCGGAGTCAG 300

355 TCCGTGTCGGATAAGCGGGACTGCGGAACTACATCACTCTCCCTGGAGAGGGCGGACTGAC 414

DB 301 GGGCTGCTGGCTGTGGCAACGCGCGGCAACCCAGCTGCTGGAACTGGTGATGGC 360
|||||

QY 415 GGGCTGCTGGCTGTGGCAACAGCCCGGACACCCAGCTGCTGGAACTGGTGATGGC 474

D5 361 ACTGTGTGACCACTTGGCGAGATGAGAGGCTATGCCCCCTTCAGTCCGGACGACAACCTAC 420

QY 4/5 ACTGTGGTGCACCTTGGCGAGATGAGAGGCTACGCCCCCCCTTCAGCCCCGACGAGAACTCC 534

D5 421 CGTGTCTGTGAGGGGACGAGGTGTATTCACCATC 459

OY 247 GGCAGCTCTCTGTGTGGTGGAGGACGTGCGAAGTCTACCTTTTGACTTCCCGAG 306
 Db 121 GGCAGAGCATCTGTGGCGACGATGATGGCTCCCAAGAGGGTCTCTGGAT 180
 OY 307 GGCAGAGCATCTGTGGCGACGATGATGGCTCCCAAGAGGGTCTCTGGAT 366
 Db 181 AAGCGGACTGCGAGACTACATCTCTCTGAGAGAGCGGAGTGGGCTGTGGCC 240
 OY 367 AAGCGGACTGCGAGACTACATCTCTCTGAGAGAGCGGAGTGGGCTGTGGCC 426
 Db 241 TGTGGACCAACAGCCCGGACCCAGCTGTGGAACCTGTGATGACACTGTGTGCA 300
 OY 427 TGTGGACCAACAGCCCGGACCCAGCTGTGGAACCTGTGATGACACTGTGTGCA 486
 Db 301 CTGGGAGATGAGAGCTACGCCCCCTTCAAGCCG 336
 OY 487 CTGGGAGATGAGAGCTACGCCCCCTTCAAGCCG 522

RESULT 8
 LOCUS AA856724 332 bp mRNA EST 09-MAR-1998
 DEFINITION od54c01.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1371744.3
 Similar to TR:064906 Q64906 SIMILAR TO GENBANK ACCESSION NUMBER
 L26081.1, mRNA sequence.
 ACCESSION AA856724
 NID 92945026
 VERSION AA856724.1 GI:2945026
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 332)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2151919.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 323.
 Location/Qualifiers
 1..332
 /organism="Homo sapiens"
 /note="vector: pT733-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, IgD-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer.
 [5'-TGTACCAATCTGAGTGGAGCGCGCTCATATTTTCTTTTCTTTT-
 3']. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT733 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /map="5: 5q23.3-5q31.1"
 /clone="IMAGE:1371744"

BASE COUNT 66 a 95 c 109 g 62 t
 ORIGIN
 Query Match 15.8%; Score 318; DB 20; Length 332;
 Best Local Similarity 99.1%; Pred. No. 0.00e+00;
 Matches 331; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

/clone.lib="NCI_CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"

Db 1 GGAATTTGGCCAGACTGAGACCGCAGACGGTCTTTTCCACGAGCAGCAGCTCTGT 60
 OY 201 GGAATTTGGCCAGACTGAGACCGCAGACGGTCTTTTCCACGAGCAGCAGCTCTGT 260
 Db 61 GTGGGTGGAGAGACTGGCAGAGTCTTACTTCTTCTTCCCGAGGCGCAGACATC 120
 OY 261 GTGGGTGGAGAGACTGGCAGAGTCTTACTTCTTCTTCCCGAGGCGCAGACATC 320
 Db 121 TGTGGCAGCGTGAATATCGGCTCCACAAGAGGCTCTGTGTGATAAGCGGAGTGGGA 180
 OY 321 TGTGGCAGCGTGAATATCGGCTCCACAAGAGGCTCTGTGTGATAAGCGGAGTGGGA 380
 Db 181 GAATACATCACTCTCTGTGAGAGCGGAGTGGGCTGTGGCTGTGGCACCACACGC 240
 OY 381 GAATACATCACTCTCTGTGAGAGCGGAGTGGGCTGTGGCTGTGGCACCACACGC 440
 Db 241 ACGGACCCCGACCTCTGTGAGAGCGGAGTGGGCTGTGGCTGTGGCACCACACGC 299
 OY 441 ACGGACCCCGACCTCTGTGAGAGCGGAGTGGGCTGTGGCTGTGGCACCACACGC 500
 Db 300 AGCTACGCCCC-TCAGCCCGGAGAGACTCC 332
 OY 501 AGCTACGCCCCCTTCAAGCCCGGAGAGACTCC 534

RESULT 9
 LOCUS AA968218 422 bp mRNA EST 19-MAY-1998
 DEFINITION uh4f09.r1 Soares mouse hypothalamus NMH9 Mus musculus cDNA clone
 IMAGE:1617929.5 similar to TR:064906 Q64906 SIMILAR TO GENBANK
 ACCESSION NUMBER L26081.1, mRNA sequence.
 ACCESSION AA968218
 NID 93142111
 VERSION AA968218.1 GI:3142111
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 422)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Getsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The Mashu-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Jan 17, 1998 this sequence version replaced gi:2043681.

Contact: Marra M/Mouse EST Project
 Mashu-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:957229
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 405.
 Location/Qualifiers
 1..422
 /organism="Mus musculus"

DEFINITION	VP796003.s1 Soares placenta Nb2HP Homo sapiens cDNA clone			
ACCESSION	U796003.1 GI:135941.3', mRNA sequence.			
NID	R33439			
VERSION	9/789297			
KEYWORDS	R33439.1 GI:789297			
SOURCE	EST.			
ORGANISM	human.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 309) Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le M., Lennon,G., Maira,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Treviski,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.			
TITLE	The WashU-Merck EST Project			
JOURNAL	Unpublished (1995)			
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wuston.wustl.edu Insert Size: 1078 High quality sequence stops: 273 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Possible reversed clone: polyT not found Insert Length: 1078 Std Error: 0.00 Seq primer: Promega -2lm13 High quality sequence stop: 273. Location/Qualifiers 1..309 /organism="Homo sapiens" /note="Organ: Placenta; Vector: pT73D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5 AACTGAGAAATTCGCGCCGCGCAGAGATTTTCTTTTCCAGACCG 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo." /db_xref="GDB:541932" /db_xref="taxon:9606" /map="10 q11.2-q22.1" /clone="IMAGE:135941" /clone_lib="Soares placenta Nb2HP" /sex="Female" /dev_stage="placenta obtained at birth (full term)" /lab_host="DH10B (ampicillin resistant)" BASE COUNT 58 a 86 c 104 g 58 t 3 others ORIGIN Query Match 14.2%; Score 286; DB 30; Length 309; Best Local Similarity 97.7%; Pred. No. 0.00e+00; Matches 295; Conservative 0; Mismatches 6; Indels 1; Gaps 1; Db 8 CCGCCAGGCGCCACTAAGAGCGGATNTANTTCCTCGCGCTTGGAAGCCATGTAG 67 QY 128 CCGCCAGGCGCCACTAAGAGCGGAGCCCGCATCTTCGCGCTTGGAAGCGCATGTAG 187 Db 68 GGCAGACCGGGTGACTTTGGCCAGACGACGACCCGACACGGTGTCTTTCCAGACCGAG 127 QY 188 GGCAGACCGGGTGACTTTGGCCAGACGACGACCCGACACGGTGTCTTTCCAGACCGAG 247 Db 128 GCAAGTCTCTGTGTGGGGGAGAGACGGGCAAGGTCTTGTGACTTCCCGGAG 187 QY 248 GCAAGTCTCTGTGTGGGGGAGAGACGGGCAAGGTCTTGTGACTTCCCGGAG 307			

LOCUS	DEFINITION	ACCESSION	NID	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	188	GCAGAAACGCATCTGTGCGCAGCGGTGATATCGGCTCCACAAAGGGTCCGTGTGATA	247									
Qy	308	GCAGAAACGCATCTGTGCGCAGCGGTGATATCGGCTCCACAAAGGGTCCGTGTGATA	367									
Db	248	AGCGGAGCTGCGAGACTACATCACTCTCTGGAGAGCGCGAGTGAAGGGCTGCTGGCC	307									
Qy	368	AGCGGAGCTGCGAGACTACATCACTCTCTGGAGAGCGCGAGTGAAGGGCTGCTGGCC	426									
Db	308	TG 309										
Qy	427	TG 428										
RESULT	12											
LOCUS	AA260340	347 bp	MRNA	EST	18-MAR-1997							
DEFINITION	va91b02.r1 Soares mouse 3ME12.5 Mus musculus cDNA clone											
	IMAGE:746763 5' similar to TR:G1000717 G1000717 SIMILAR TO GENBANK											
	ACCESSION NUMBER L26081. ; mRNA sequence.											
ACCESSION	AA260340											
NID	g1896824											
VERSION	AA260340.1	GI:1896824										
KEYWORDS	EST.											
SOURCE	house mouse.											
ORGANISM	Mus musculus											
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;											
AUTHORS	Eukarya; Rodentia; Sciurognathi; Muridae; Murinae; Mus.											
	1 (bases 1 to 347)											
	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,											
	Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,											
	Schellenberg,R., Steptoe,M., Tan,F., Underwood,K., Moore,B.,											
	Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and											
	Waterston,R.											
TITLE	The WashU-HMI Mouse EST Project											
JOURNAL	Unpublished (1996)											
COMMENT	On Sep 12, 1996 this sequence version replaced gi:1404539.											
	Contact: Marra M/Mouse EST Project											
	WashU-HMI Mouse EST Project											
	Washington University School of Medicine											
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108											
	Tel: 314 286 1800											
	Fax: 314 286 1810											
	Email: mouseest@watson.wustl.edu											
	This clone is available royalty-free through LNL ; contact the											
	IMAGE Consortium (info@image.lln1.gov) for further information.											
	MGI:455747											
	Possible reversed clone; similarity on wrong strand											
	Seq primer: -28m13 rev2 ET from Amersham											
	High quality sequence stop: 296.											
FEATURES	Location/Qualifiers											
source	1..347											
	/organism="Mus musculus"											
	/strain="C57BL/6J"											

Query Match 13.3%; Score 267; DB 10; Length 347;
 Best Local Similarity 88.7%; Pred. No. 0.00e+00;
 Matches 306; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Db 1 GAGAGACACACCTGACACAGACCCGAGGCTCTCTCATAGTGTCCCGAGTACCCAGT 60
 |||||
 QY 734 GAGAGACACACCTGACACAGATCTGAGGCTCTCTCATAGTGTCCCGAGTACCCAGT 793
 |||||
 Db 61 TGTCAGAGGGGACAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
 |||||
 QY 794 TGTCAGAGGGGACAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 853
 |||||
 Db 121 TGAAGGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
 |||||
 QY 854 TGAAGGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 913
 |||||
 Db 181 ATGTCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 240
 |||||
 QY 914 ACGTCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 973
 |||||
 Db 241 TCTCCACCCCTGAGACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 |||||
 QY 974 TCTCCACCCCTGAGACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1033
 |||||
 Db 301 TCTTCCCTACTCATCGCTCAAGGCTACACATGAGGCTTTTCCA 345
 |||||
 QY 1034 TCTTCCCTACTCATCGCTCAAGGCTACACATGAGGCTTTTCCA 1078
 |||||

RESULT 13
 LOCUS AA747293 157 bp mRNA EST 17-MAR-1999
 DEFINITION nx88a10.s1 NCI-CGAP-GCBI Homo sapiens cDNA clone IMAGE:1269306 3'
 similar to TR:064906 064906 SIMILAR TO GENBANK ACCESSION NUMBER
 L26081.1, mRNA sequence.
 ACCESSION AA747293
 NID 92787251
 VERSION AA747293.1 GI:2787251
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 157)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:693526.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bdrp/image/image.html

Insert Length: 825 Std Error: 0.00
 Seq primer: -40ml3 fwd. RT from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

source 1. 157

/organism="Homo sapiens"
 /note="Vector: pT7D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, 19D-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 [5'-TGTTCAATCTGAGTGGAGCGGCGGCTCATTTTCTTTTCTT-
 3']. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7D vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /map="1p: 5"
 /clone="IMAGE:1269306"
 /clone_lib="NCI-CGAP-GCBI"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"

BASE COUNT 33 a 43 c 54 g 26 t 1 others
 ORIGIN

Query Match 6.4%; Score 129; DB 17; Length 157;
 Best Local Similarity 94.6%; Pred. No. 3.51e-181;
 Matches 141; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
 Db 5 GGCAGAGACGATCTGTGCGCAGCGTGAATATGCTCCACAAAGGCTCTGTGTGAT 64
 |||||
 QY 307 GGCAGAGACGATCTGTGCGCAGCGTGAATATGCTCCACAAAGGCTCTGTGTGAT 366
 |||||
 Db 65 AAGCGGAGCTGCGAGACTACATCATCTCTCTGTGAGAGGGGAGTGGGCTGTGGC 124
 |||||
 QY 367 AAGCGGAGCTGCGAGACTACATCATCTCTCTGTGAGAGGGGAGTGGGCTGTGGC 426
 |||||

RESULT 14
 LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
 DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPLI cDNA library Oryza sativa
 cDNA clone 97SN1787, mRNA sequence.
 ACCESSION AA754459
 NID 92801165
 VERSION AA754459.1 GI:2801165
 KEYWORDS EST.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
 Poaceae; Oryza.
 REFERENCE 1 (bases 1 to 252)
 Naim, B.H., Kim, W.Y., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P.,
 Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y.,
 Lee, M.C., and Eun, M.Y.
 Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
 Unpublished (1998)
 JOURNAL On Jan 14, 1998 this sequence version replaced gi:1797457.
 COMMENT

Contact: Eun M. Y.
 Department of Cytogenetics
 National Inst. of Agri. Sci. and Tech, RDA
 Suwon, Kyunggi-do, Korea
 Tel: 82 331 290 0301
 Fax: 82 331 290 0307
 Email: myeun@suno2.asti.re.kr
 Submitted by Beek Hie Naim, Dept of Biological Science, Myongji
 University, Yongin, Korea. 449-728 bnamh@loserver.myongji.ac.kr
 Seq primer: M13 Reverse Primer.
 Location/Qualifiers

FEATURES

source 1. 252

/organism="Oryza sativa"
 /cultivar="Milyang23"
 /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
 XhoI; Directional cDNA library inserted into lambda ZAPLI
 vector at 5' end with EcoRI and 3' end with Xho I site."
 /db_xref="taxon:4530"

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

```
Run on:      Wed Aug  4 01:06:00 1999;  MasPar time 1236.57 Seconds
```

Tabular output not generated.

```

Title: >US-09-240-410-3
Description: (1-712) from US09240410.seq
Perfect Score: 708
N.A. Sequence:
Comp: 1 CCGCCGCGCCCAAGGCCA.....CCAGGTCGGGAAAGTTCAN 712
      GGGGACGGCGGGATCCCGGT.....GATCCACACCTTTCAGTN

```

Scoring table: TABLE default

```
Nmatch  STD : Dbase 0; Query 0
```

Searched: 2883791 seqs, 1171580779 bases x 2

Post-processing: Minimum Match 0%

Database: emb1-est58

Database: genbank-est111

Statistics: Mean 10.492; Variance 1.935; scale 5.421

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description	Pred.	No.
1	393	55.5	476	18	AA827795		od08a05.s1	NC1_CGAP_GC	0.00e+00	
2	390	55.1	459	24	A1206011		qg14h09.x1	Soares_plac	0.00e+00	
3	372	53.5	446	30	H02902		y139f01.s1	Soares_plac	0.00e+00	
4	379	52.5	469	32	A0395532		o117c05.s1	Soares_sene	0.00e+00	
5	356	50.3	424	34	W47265		zc39a12.r1	Soares_sene	0.00e+00	
6	309	43.6	336	17	AA731451		n294g05.s1	NC1_CGAP_GC	0.00e+00	
7	7	42.2	309	30	R33439		yh19d03.s1	Soares_plac	0.00e+00	
8	296	41.8	332	20	AA856724		od54c01.s1	NC1_CGAP_GC	0.00e+00	
9	129	18.2	157	17	AA747293		nx88a10.s1	NC1_CGAP_GC	2.48e+213	
10	75	10.6	347	10	AA260340		va31b02.r1	Soares_mous	2.51e+99	

C	11	57	8.1	252.17	AA754459	97SN1787	Rice	Immature	1.39e-61
C	12	53	7.5	252.17	AA754459	97SN1787	Rice	Immature	6.33e-56
C	13	51	7.2	247.17	AA734458	97SN1784	Rice	Immature	3.81e-52
C	17	48	6.8	247.17	AA754458	97SN1784	Rice	Immature	1.51e-46
C	15	41	5.8	344.30	R33537	yh9p0c3	rt Soares plac	7.18e-34	4
C	16	30	4.2	2275.20	AF034173	AF034173	Human mRNA (T	1.70e-15	1
C	17	30	4.2	2275.20	AF034173	AF034173	Human mRNA (T	1.70e-15	1
C	18	27	3.8	492.35	AA023538	mh5d07.r1	Soares mous	5.77e-11	1
C	19	24	3.4	369.23	AI145074	UI-R-BT0-pl-r-05-0-U-I-	1.00e-06	1	0
C	20	24	3.4	441.13	AA466386	v96b0b3.r1	Barstead mo	1.00e-06	1
C	21	24	3.4	447.17	AA405549	LD15364.5prtime	LD Dros	1.00e-06	1
C	22	24	3.4	489.38	AA072554	HS_2247_x2	AI10 MR CIT	1.00e-06	1
C	22	24	3.4	543.28	AI544302	SD10805.5prtime	SD Dros	1.00e-06	1
C	23	24	3.4	627.24	AI258321	LP01477.5prtime	LP Dros	1.00e-06	1
C	25	24	3.4	751.21	AA951002	LD31216.5prtime	LD Dros	1.00e-06	1
C	26	23	3.2	377.26	AA819595	UI-R-A0-bt-h-10-0-U-I.s	2.16e-05	2	1
C	27	23	3.2	415.24	AI233144	ESR1228832	Normalized r	2.16e-05	1
C	28	23	3.2	437.24	AI222186	gh034o1.x1	Soares_MFL	2.16e-05	1
C	29	23	3.2	465.24	AI228822	ESR1225517	Normalized r	2.16e-05	1
C	30	23	3.2	480.26	AA900201	UI-R-A0-bt-h-10-0-U-I.s	2.16e-05	2	1
C	31	23	3.2	490.26	AI365650	ap20h08.x1	Schiller ol	2.16e-05	1
C	32	23	3.2	498.9	AA190645	zg44a06.r1	Stratogene	2.16e-05	1
C	33	23	3.2	503.24	AI227673	ESR1224368	Normalized r	2.16e-05	1
C	34	23	3.2	515.26	N22864	yx56f09.s1	Soares meda	2.16e-05	1
C	35	23	3.2	522.26	AI406764	ESR1235051	Normalized r	2.16e-05	1
C	36	23	3.2	525.23	AI171412	ESR1217372	Normalized r	2.16e-05	1
C	37	23	3.2	544.23	AI105159	ESR1214448	Normalized r	2.16e-05	1
C	38	23	3.2	638.23	AI105448	ESR1214737	Normalized r	2.16e-05	1
C	39	23	3.2	660.20	AF034177	AF034177	Human mRNA (T	2.16e-05	1
C	40	23	3.2	660.20	AF034177	AF034177	Human mRNA (T	2.16e-05	1
C	41	23	3.2	686.23	AI172048	ESR1218043	Normalized r	2.16e-05	1
C	42	23	3.2	714.37	AF001895	Homo sapiens genomic	D	4.23e-04	4
C	43	22	3.1	77.30	R82156	6E3 Chromosome 21	exon	4.23e-04	4
C	44	22	3.1	442.41	AF0322415	RPC111-92N3.TJ	RPC111	4.23e-04	4
C	45	22	3.1	1287.20	AF038250	AF038250	Human mRNA (T	4.23e-04	4

ALIGNMENTS

RESULT	1
LOCUS	AA827795 476 bp mRNA EST 25-MAR-1998
DEFINITION	cd088405.s1 NCI-CGAP-CG81 Homo sapiens cDNA clone IMAGE:1367312 3 similar to TR:Q64906 Q64906 SIMILAR TO GENBANK ACCESSION NUMBER L26081. ; mRNA sequence.
ACCESSION	AA827795
NID	Q2900158
VERSION	AA827795.1 GI:2900158
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 476).
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	On Jan 14, 1998 this sequence version replaced gi:1877580.

Contact: Robert Strausberg, Ph.D.
 Tel.: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdir/image/image.html
 Insert Length: 1737 Std Error: 0.00

OY 367 CTGGCGAGAGTGGAGCTACGCCCCCTTACGCCGAGAGAGCTCCGTCTGT 426

Db 431 TT-GAAGGGGAGAGGTGTATTCACCATC 459
|||||
OY 427 TTGAAGGGGAGAGGTATTCACCATC 456

RESULT 3
LOCUS H02902 446 bp mRNA EST 20-JUN-1995
DEFINITION Y139f01.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
IMAGE:151129 3', mRNA sequence.

ACCESSION H02902
NID 9865835
VERSION H02902.1 GI:865835
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 446)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Paterson, J., Rifkin, B., Rohlfing, T., Soares, M., Tan, F., and
Trevaskis, E., Waterston, R., Williamson, A., Woldmann, P., and
Wilson, R.
TITLE The Mashu-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 1575
High quality sequence stops: 196
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1575 Std Error: 0.00
Seq primer: Promega -2Lm13
High quality sequence stop: 196.
Location/Qualifiers
1. .446
/organism="Homo sapiens"
/note="Organ: placenta; Vector: p773D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
AACTGGAAGAATTCGCGCCGAGAGATTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo.
/db_xref="GDB:563192"
/db_xref="taxon:9606"
/map="851H08:12:12q24.13-12q24.31"
/clone="IMAGE:151129"
/clone_lib="Soares placenta Nb2HP"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 93 a 115 c 136 g 93 t 9 others
ORIGIN

Query Match 53.5%; Score 379; DB 30; Length 446;
Best Local Similarity 92.9%; Pred. No. 0.00e+00;
Matches 417; Conservative 0; Mismatches 29; Indels 3; Gaps 2;
Db 1 GCCGCTGGAAGGCGCATGTAGGCGAGACCGGCTTGGCCACAGTGGCGGAC 60
OY 46 GCCGCTGGAAGGCGCATGTAGGCGAGACCGGCTTGGCCACAGTGGCGGAC 105

Db 61 ACGGTCTTTTCACGAGCCAGGACGCTCTCTGTGTGGGTGGAGACGTGGCAAGCTC 120
|||||
OY 106 ACGGTCTTTTCACGAGCCAGGACGCTCTCTGTGTGGGTGGAGACGTGGCAAGCTC 165
|||||

Db 121 TACCTCTTGAATCTCCCGAGGCGCAAGAACCATCTGTGCCGACCGTGAATATCGGCTCC 180
|||||
OY 166 TACCTCTTGAATCTCCCGAGGCGCAAGAACCATCTGTGCCGACCGTGAATATCGGCTCC 225
|||||

Db 181 ACAAGGGGCTCTGTCTGATTAAGCGGAGATCGGAGACTACATCTCTCTGGAGAG 240
|||||
OY 226 ACAAGGGGCTCTGTCTGATTAAGCGGAGATCGGAGACTACATCTCTCTGGAGAG 285
|||||

Db 241 CGGANTGAGGGGCTGCTGGCTGTGG-ACCAAGCCCGGCAACCCAGCTGTGAAACCTG 299
|||||
OY 286 CGGAGTGAAGGGGCTGCTGGCTGTGG-ACCAAGCCCGGCAACCCAGCTGTGAAACCTG 345
|||||

Db 300 GGTGATNGCACATNTGCTACCTTGGCGAGATGAGAGCTACGCCCTTTACGCCNAGC 359
|||||
OY 346 GTGAATGCACCTGTGTGCTGCACCTTGGCGAGGTGAGAGCTACGCCCTTTACGCCNAGC 405
|||||

Db 360 GAGACTTCC--TGCTCTNTTTTAAAGGGGAGCAAGTTATTCANCATCCNGAAGCA 417
|||||
OY 406 GAGACGCTCCCGTGTCTGTGTTGAAGGGGAGCAAGTGTATTCACCATCCGGAAGCA 465
|||||

Db 418 AAGATTTACAATTGGGAAGATCTTCGTT 446
|||||
OY 466 AAGATTTACAATTGGGAAGATCTTCGTT 494
|||||

RESULT 4
LOCUS A1095332 469 bp mRNA EST 24-SEP-1998
DEFINITION OY17C05.s1 Soares senescent fibroblasts Nb2SF Homo sapiens cDNA
clone IMAGE:1666088 3' similar to TR:064906 064906 SIMILAR TO
GENBANK ACCESSION NUMBER U26081.; mRNA sequence.
ACCESSION A1095332
NID 93434308
VERSION A1095332.1 GI:3434308
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 469)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/nciccap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2286378.

FEATURES
source
1. .469
/organism="Homo sapiens"
/note="Vector: p773D (Pharmacia) with a modified
polylinker V_type: phagemid; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5',
TGTTCACATCTGAAGTGGAGGCGCGCATTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo."
/db_xref="taxon:9606"
/map="13"

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 336)
NCL-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Nov 29, 1993 this sequence version replaced gi.430441.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.liml.gov/db/ftp/image/image.html

Insert Length: 1045 Std Error: 0.00
Seq primer: -40m13 fwd. E7 from Amersham
High quality sequence stop: 202.

FEATURES
Location/Qualifiers
1..336
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI). Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAGTGGAGGCGCCGCTCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1303160"
/clone_lib="NCL-CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
BASE COUNT 66 a 95 c 113 g 62 t

ORIGIN
Query Match 43.6%; Score 309; DB 17; Length 336;
Best Local Similarity 97.3%; Pred. No. 0.00e+00;
Matches 328; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

Db 1 GGGCGAGACCGGGTGGACCTTGGCCAGACTGAGCGGACAGCGGTCTTTCACGAGCA 60
Qy 67 GGGCGAGACCGGGTGGACCTTGGCCAGACTGAGCGGACAGCGGTCTTTCACGAGCA 126
Db 61 GGCACCTCCTCTGTGTGGGTGGAGAGCTGCAAGTCTTACTCTTTGACTTCCCGAG 120
Qy 127 GGCACCTCCTCTGTGTGGGTGGAGAGCTGCAAGTCTTACTCTTTGACTTCCCGAG 186
Db 121 GGCAGAAAGCATCTGTGGCGACGCTGAATATCGGCTCCCAAAGGGTCTGTGTGAT 180
Qy 187 GGCAGAAAGCATCTGTGGCGACGCTGAATATCGGCTCCCAAAGGGTCTGTGTGAT 246
Db 181 AAGCGGAGCTGGAGAACATCACTCTCCGAGAGAGCGGAGAGGGGCTGTGGGCC 240
Qy 247 AAGCGGAGCTGGAGAACATCACTCTCCGAGAGAGCGGAGAGGGGCTGTGGGCC 306
Db 241 TGTGCGACCAAGCGCCGAGCGGACGCTGGAACCTTGTAATGGCACTGTGTGCA 300
|||||

Qy 307 TGTGCGACCAAGCGCCGAGCCGACCCGACGCTGTGAACTGTGAATG-CACGTGTGTGCCA 365
Db 301 C-TTGGCGAGATGAGAGCTACGCCCTTCAGCGC 336
Qy 366 CTTGGCGAGATGAGAGCTACGCCCTTCAGCGCG 402
|||||

RESULT 7
LOCUS R33439 309 bp mRNA EST 28-APR-1995
DEFINITION yb79d03.s1 Soares placenta NB2HP Homo sapiens cDNA clone
IMAGE:135941.3, mRNA sequence.
ACCESSION R33439
NID 9789297
VERSION R33439.1 GI:789297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 309)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hulman, M., Kucaba, T., Le M., Lennon, G., Marra, M.,
Patterson, D., Rifkin, B., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
TITLE The Mashu-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1078
High quality sequence stops: 273
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.liml.gov) for further information.
Possible reversed clone: polyT not found
Insert Length: 1078 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 273.

FEATURES
Location/Qualifiers
1..309
/organism="Homo sapiens"
/note="Organ: placenta; Vector: pT73D (Pharmacia) with a
modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
ACTGGAAGAATTGCGCGCCGACAGATTTTCTTTTCTTTTCTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, library
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="GDB:541932"
/db_xref="taxon:9606"
/map="10 q11.2-q22.1"
/clone="IMAGE:135941"
/clone_lib="Soares placenta NB2HP"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 58 a 86 c 104 g 58 t 3 others

ORIGIN
Query Match 42.2%; Score 299; DB 30; Length 309;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 CCGCTGCGGCCGACGAGCGGACCTTAAGAGGCGATNCTTTCGCGCTGCAAGGC 60
Qy 1 CCGCTGCGGCCGACGAGCGGACCTTAAGAGCGGATNCTTTCGCGCTGCAAGGC 60
|||||

Db 61 CATGTAGGCGACGACGGGTGACTTTGGCCAGACTGAGCCGACACGGTCTTTTCAC 120
|||||
Qy 61 CATGTAGGCGACGACGGGTGACTTTGGCCAGACTGAGCCGACACGGTCTTTTCAC 120
121 GAGCCAGGACAGCTCTCTGTGTGGGTGGAGAGAGCTGGCAAGTCTACTCTTTGACTTC 180
|||||
Qy 121 GAGCCAGGACAGCTCTCTGTGTGGGTGGAGAGAGCTGGCAAGTCTACTCTTTGACTTC 180
121 GAGCCAGGACAGCTCTCTGTGTGGGTGGAGAGAGCTGGCAAGTCTACTCTTTGACTTC 180
181 CCCAGGCGCAGAACGATCTGTGCGCAGCGGTGATATCGCTCCACAAAGGGTCTCTGT 240
Qy 181 CCCAGGCGCAGAACGATCTGTGCGCAGCGGTGATATCGCTCCACAAAGGGTCTCTGT 240
241 CTGATTAAGCGGACGCTGCGAGAACTACATCTCTCTCTGAGAGAGCGGAGTGAAGGAGCT 300
|||||
Qy 241 CTGATTAAGCGGACGCTGCGAGAACTACATCTCTCTCTGAGAGAGCGGAGTGAAGGAGCT 300
Db 301 GCTGGCCTG 309
|||||
Qy 301 GCTGGCCTG 309
300 GCTGGCCTG 308

RESULT 8
LOCUS AA56724 332 bp mRNA EST 09-MAR-1998
DEFINITION o54601.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1371744 3'
similar to TR:064906 064906 SIMILAR TO GENBANK ACCESSION NUMBER
126081. ; mRNA sequence.
AA56724
G2945026
VERSION AA56724.1 GI:2945026
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 332)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 19, 1998 this sequence version replaced gi:2151919.

JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html

FEATURES
SOURCE
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 323.
Location/Qualifiers
1..332
/organism="Homo sapiens"
/note="Vector: p7713D-PAC (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCATCTGAAGTGGAGAGCGGCCCTCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7713 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 66 a 95 c 109 g 62 t
ORIGIN
Query Match 41.8%; Score 296; DB 20; Length 332;
Best Local Similarity 97.6%; Pred. No. 0.00e+00;
Matches 324; Conservative 0; Mismatches 4; Indels 4; Gaps 4;
/db_xref="taxon:9606"
/map="5: 5623.3-5631.1"
/clone="IMAGE:1371744"
/clone_11b="NCI-CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

Db 1 GGACTTTGGCGACAGCTGAGCCGACAGCGGTGCTTTTCACAGACCCAGGAGCTCTGT 60
|||||
Qy 81 GGACTTTGGCGACAGCTGAGCCGACAGCGGTGCTTTTCACAGACCCAGGAGCTCTGT 140
61 GTGGGTGGAGAGAGCTGGCGAAGGCTACCTCTTTGACTTCCCGAGGCGAAGACGATC 120
Qy 141 GTGGGTGGAGAGAGCTGGCGAAGGCTACCTCTTTGACTTCCCGAGGCGAAGACGATC 200
Db 121 TGTGCGCAGCGTGAATTCGCTCCACAAAGGGTCTCTGTGATTAAGCGGAGCTGCGA 180
|||||
Qy 201 TGTGCGCAGCGTGAATTCGCTCCACAAAGGGTCTCTGTGATTAAGCGGAGCTGCGA 260
Db 181 GAACATCATCTCTCTCTGAGAGAGGAGTGAAGGAGGCTGCTGCTGAGCAGCAACGC 240
|||||
Qy 261 GAACATCATCTCTCTCTGAGAGAGGAGTGAAGGAGGCTGCTGCTGAGCAGCAACGC 320
Db 241 ACGGACCCGAGCTGCTGAGAACCTGTGATGAGCTGTGTGTC-AC-TTGGCAGATGA 298
|||||
Qy 321 CCGGACACCCAGCTGCTGAGAACCTGTGATG-CACTGTGTGTCACACTTGGGAGAGTG 379
Db 299 GAGGCTACGCCC-TTCAAGCCCGGACGAGAAC 329
|||||
Qy 380 GAGGCTACGCCCCCTTCAGCCCGGACGAGAAC 411

RESULT 9
LOCUS AA747293 157 bp mRNA EST 17-MAR-1999
DEFINITION nx88a10.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1269306 3'
similar to TR:064906 064906 SIMILAR TO GENBANK ACCESSION NUMBER
126081. ; mRNA sequence.
AA747293
G2787251
VERSION AA747293.1 GI:2787251
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 157)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Apr 14, 1993 this sequence version replaced gi:693526.

JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html

Insert Length: 825 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham

FEATURES High quality sequence stop: 1.
Location/Qualifiers
1. 157
source

/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, 19D-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CEBR). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCATCTGTAAGTGGAGCGCGCTCATTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 33 a 43 c 54 g 26 t 1 others
ORIGIN
/map="1p: 5"
/clone="IMAGE:1269306"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

Query Match 18.2%; Score 129; DB 17; Length 157;
Best Local Similarity 94.6%; Pred. No. 2,48e-213;
Matches 141; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Db 5 GGCAGAGACGATCTGTGGCAGCGTGAATCGCTCCACAAAGGGGTCTGTGGAT 64
|||||
Qy 187 GGCAGAGACGATCTGTGGCAGCGTGAATCGCTCCACAAAGGGGTCTGTGGAT 246
|||||
Db 65 AAGCGAGCTGCGAAGATCATCTCTCTGGAGAGCGCGAGTGAAGGGGCTGTGGC 124
|||||
Qy 247 AAGCGAGCTGCGAAGATCATCTCTCTGGAGAGCGCGAGTGAAGGGGCTGTGGC 306
|||||
Db 125 TGTGG-ACCAAGCGCGCGACCCAGCTG 152
|||||
Qy 307 TGTGGACCAAGCGCGCGACCCAGCTG 335
|||||

RESULT 10
LOCUS AA260340 347 bp mRNA EST 18-MAR-1997
DEFINITION va91002.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone
IMAGE:746763 5' similar to TR:G1000717 G1000717 SIMILAR TO GENBANK
ACCESSION NUMBER U26081. ; mRNA sequence.

ACCESSION AA260340
NID G1896824
VERSION AA260340.1 GI:1896824
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 347)
Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Scheibler, R., Stepien, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HMT Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1404539.

Contact: Maria M/Mouse EST Project
WashU-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

FEATURES
source

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:455747
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 296.
Location/Qualifiers
1. 347
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: whole fetus; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker: Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCATCTGTAAGTGGAGCGCGCTCATTTTCTTTT-
3'], on total mouse RNA [provided by Minoru Ko, Wayne
State Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 75 a 102 c 90 g 80 t
ORIGIN
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
/clone="unknown"
/tissue_type="fetus"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"

Query Match 10.6%; Score 75; DB 10; Length 347;
Best Local Similarity 91.2%; Pred. No. 2.51e-99;
Matches 83; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 1 GAGAGCAACACCTGACAGACCCGAGGCTCTCAATGTGTCCGAGTACCCAGT 60
|||||
Qy 620 GAGAGCAACACCTGACAGACCCGAGGCTCTCTCAATGTGTCCGAGTACCCAGT 679
|||||
Db 61 TGTGAGGGGGGAGACAGGCTGTGAGAGTTC 91
|||||
Qy 680 TGTGAGGGGGGAGACAGGCTGTGAGAGTTC 710
|||||

RESULT 11
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPIT cDNA library Oryza sativa
cDNA clone 97SN1787, mRNA sequence.

ACCESSION AA754459
NID 92801165
VERSION AA754459.1 GI:2801165
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Oryza.
AUTHORS 1 (bases 1 to 252)
Nahm, B.H., Kim, J.K., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P.,
Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y.,
Lee, M.C. and Eun, M.Y.

TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1797457.

Contact: Eun M.Y.
Department of Cyto genetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggi-do, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@suno2.astl.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhnahe@bioserver.myongji.ac.kr

FEATURES
source

Seq primer: M13 Reverse Primer.

Location/Qualifiers

1..252

/organism="Oryza sativa"

/cultivar="Milyang23"

/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:

XhoI; Directional cDNA library inserted into lambda ZAPII

vector at 5' end with EcoRI and 3' end with Xho I site."

/db_xref="taxon:4530"

/map="6"

/clone="97SN1787"

/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"

/tissue_type="Immature Seed"

/dev_stage="5 days after pollination"

/lab_host="E. coli SOLR"

BASE COUNT 5 a 21 c 12 g 35 t 179 others

ORIGIN

Query Match 8.1%; Score 57; DB 17; Length 252;

Best Local Similarity 9.6%; Pred. No. 1,39e-63;

Matches 22; Conservative 122; Mismatches 80; Indels 4; Gaps 4;

Db 21 TSYBCHGNBVMVCVASHGNMVSVNCTBRGTHCDCKNVNMTWTGTVNMBNVSQDMHY 79

Cp 526 TGTACAGCTCACTCTCCGCCCGGATCGCGGACGACGAGATCTCCCAATGTAATTCG 467

Db 80 MBVBNKVDVGNHTRCSMRBVTBMAYHYDYNCBVNNNDYHMMHBMVBTCMTCTM 139

Cp 466 TTGCTTCCGGATGCTGGAAGGGGGGCTGACCTCCA-CCTGCGCCAAAGT-GGACACAGATGCAT 408

Db 140 MCBMHNKTKTASGWHSTNDVKSSTNTWGTBSYDKSMHGYCSBBVKYHTVSTRA 199

Cp 407 TCGTCCGGGCTGAAGGGGGGCTGACCTCCA-CCTGCGCCAAAGT-GGACACAGATGCAT 350

Db 200 TRSYTCVARKVMMTKKVVKKYHVBGCHBTDSKCKTMNTKNTKHYM 247

Cp 349 TCACAGAGTTCACAGAGCTGGGGGCTGTGTCGACACAGCCCA 302

RESULT 12

LOCUS AA754459 252 bp mRNA EST 20-JAN-1998

DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa

ACCESSION AA754459

NID 92801165

VERSION AA754459.1 GI:2801165

KEYWORDS EST.

SOURCE Oryza sativa.

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;

Poaceae; Oryza.

REFERENCE 1 (bases 1 to 252)

Authors Namh,B.H., Kim,W.Y., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,

Lee,M.C. and Eun,M.Y.

Large-scale sequencing Analysis of ESTs from Rice Immature Seed

Unpublished (1998)

On Jan 14, 1998 this sequence version replaced gi:1797457.

TITLE

JOURNAL

COMMENT

Contact: Eun M.Y.

Department of Cyto genetics

National Inst. of Agri. Sci. and Tech, RDA

Suwon, Kyungjido, Korea

Tel: 82 331 290 0301

Fax: 82 331 290 0307

Email: myeunesun20.astl.re.kr

Submitted by Baek Hie Namh, Dept of Biological Science, Myongji

University, Yongin, Korea, 449-728 bhnahm@bserver.myongji.ac.kr

Seq primer: M13 Reverse Primer.

FEATURES

source

1..252

/organism="Oryza sativa"

/cultivar="Milyang23"

/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:

XhoI; Directional cDNA library inserted into lambda ZAPII

vector at 5' end with EcoRI and 3' end with Xho I site."

/db_xref="taxon:4530"

/map="6"

/clone="97SN1787"

/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"

/tissue_type="Immature Seed"

/dev_stage="5 days after pollination"

/lab_host="E. coli SOLR"

BASE COUNT 5 a 21 c 12 g 35 t 179 others

ORIGIN

Query Match 7.5%; Score 53; DB 17; Length 252;

Best Local Similarity 12.9%; Pred. No. 6.33e-56;

Matches 30; Conservative 114; Mismatches 84; Indels 4; Gaps 4;

Db 21 TSYBCHGNBVMVCVASHGNMVSVNCTBRGTHCDCKNVNWS-TMTGTVNMBNVSQDMHY 79

Qy 302 TGGCTTGCGACACACAGCGCCGACCCAGCTGCTGGAACCTGATGCATGCTGCT 361

Db 80 MBV-BNTRVDVGNHTRCSMRBVTBMAYHYDYNCBVNNNDYHMMHBMVBTCMTCT 138

Qy 362 GCCACTTGCGGAGAGTGAGGCTACGCCCTTCAGCCCGG-ACGAGACGCTCCCGTGG 420

Db 139 MNCMBHNKTKTASGWHSTNDVKSSTNTWGTBSYDKSMHGYCSBBVKYHTVSTRA 198

Qy 421 TTCTGTTTGAAGGAGACCAAGTATTCACCACTCCGGAAGCAAGATATCAATTGG 480

Db 199 ATRSYTCVARKVMMTKKVVKKYHVBGCHBTDSKCKTMNTKNTKHYM 250

Qy 481 GAAGATCC-TGCGTTCGCCGCGATCGGGGCGAGAGTACGCTGACACCAAGT 531

RESULT 13

LOCUS AA754458 247 bp mRNA EST 20-JAN-1998

DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa

CDNA clone 97SN1784, mRNA sequence.

ACCESSION AA754458

NID 92801164

VERSION AA754458.1 GI:2801164

KEYWORDS EST.

SOURCE Oryza sativa.

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;

Poaceae; Oryza.

REFERENCE 1 (bases 1 to 247)

Authors Namh,B.H., Kim,W.Y., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,

Lee,M.C. and Eun,M.Y.

Large-scale sequencing Analysis of ESTs from Rice Immature Seed

Unpublished (1998)

On Jan 14, 1998 this sequence version replaced gi:1797455.

TITLE

JOURNAL

COMMENT

Contact: Eun M.Y.

Department of Cyto genetics

National Inst. of Agri. Sci. and Tech, RDA

Suwon, Kyungjido, Korea

Tel: 82 331 290 0301

Fax: 82 331 290 0307

Email: myeunesun20.astl.re.kr

Submitted by Baek Hie Namh, Dept of Biological Science, Myongji

University, Yongin, Korea, 449-728 bhnahm@bserver.myongji.ac.kr

Seq primer: M13 Reverse Primer.

FEATURES

source

1..247

/organism="Oryza sativa"

/cultivar="Milyang23"

/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:

XhoI; Directional cDNA library inserted into lambda ZAPII

vector at 5' end with EcoRI and 3' end with Xho I site."

```
/db_xref="taxon:4530"
/map="6"
/clone_lib="Rice Immature Seed Lambda ZAP11 CDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"

BASE COUNT      7 a      16 c      21 g      34 t      169 others
ORIGIN

Query Match      7.2%; Score 51; DB 17; Length 247;
Best Local Similarity 13.4%; Pred. No. 3.81e-52;
Matches 33; Conservative 121; Mismatches 87; Indels 6; Gaps 6;

Db 1 HMDCTNNVWRGCCCBAMNKHHTHMTBMCVRRVGTNNKHNHGRRTTNDCSDNA 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 686 CTGCACACTGGGCGACCGGACACATTAG-AGGAGCCTCAGATCTTGACAGATT 628
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 HCRYTBWYARSKYGYGTBYSMNVDTNGTGKTTVNVHSGMNNRCSNVYVWBT 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 627 GTCTCTCGAAGAGTA-GTAGATCTTGCAI-CGTAAAGCCTGCTTGTCACAGATG 570
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 AYCDYHYBDRANHVDTFCTNDRGYCNVTASDNG-TSATKRYTGDKTSDCGGCKWK 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 569 GTGGCTTTGTAAGTGTGCTGATGACAGTACAGTGTGTAACAGCTCAGCTCTCG 510
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 VTGSSBYBRCGVNVRTSMVTDKSTKBSMDMSRVRVHGRMMBNKKRMSRWTD 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 509 CCC-CGGATCGGGGACCGAGAGCTTCCCAATTGT-NAATCTTGCTTCCGGATGG 452
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 TKTWTR 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 451 TGGATA 445

RESULT 14
LOCUS      AA754458      247 bp      mRNA      EST      20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAP11 CDNA library Oryza sativa
            CDNA clone 97SN1784, mRNA sequence.
ACCESSION  AA754458
NID        92801164
VERSION    AA754458.1 GI:2801164
KEYWORDS   EST.
SOURCE     Oryza sativa.
ORGANISM   Oryza sativa.
REFERENCE  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
            Poaceae; Oryza.
            1 (bases 1 to 247)
            Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
            Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,J.I., Kang,K.Y.,
            Lee,M.C., and Eun,M.Y.
            Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
            Unpublished (1998)
            On Jan 14, 1998 this sequence version replaced gi:1797455.

TITLE
JOURNAL
COMMENT

Contact: Eun M.Y.
Department of Cyto genetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggi-do, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea, 449-728 bhnamh@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers
1..247
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional CDNA library inserted into lambda ZAP11
vector at 5' end with EcoRI and 3' end with Xho I site."
```

```
/db_xref="taxon:4530"
/map="6"
/clone_lib="Rice Immature Seed Lambda ZAP11 CDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"

BASE COUNT      7 a      16 c      21 g      34 t      169 others
ORIGIN

Query Match      6.8%; Score 48; DB 17; Length 247;
Best Local Similarity 12.4%; Pred. No. 1.51e-46;
Matches 28; Conservative 103; Mismatches 93; Indels 2; Gaps 2;

Db 6 MNTVWGGCCCBAMNKHHTHMTBMCVRRVGTNNKHNHGRRTTNDCSDNACRYT 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 21 CTTAAGACGGGATGCTNCTTANNTCTTCCGCTGGAAGGCATGTAGGACGAGCCGGT 80
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 VBWYARSKYGYGTBYSMNVDTNGTGKTTVNVHSGMNNRCSNVYVWBTAYCDY 125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 81 GGACTTGGCCAGACT-GAGCCGACACAGGTGCTTTTCACAGCAGCAGCTCTCTG 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 126 BHVBRANHVDTFCTNDRGYCNVTASDNGTATKRVGYDKTSDCGGCKWKRYGSS 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 140 TGTGGGTGGAGAGAGCTGCAAGTCTACCTCTTGAC-TTCCCGGAGGCAAGACGA 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 186 VBVRGCVNVRTSMVTDKSTKBSMDMSRVRVHGRMMBNKKR 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 199 TGTGCGCAGGATGTAATCGGCTCCACAAAGGGCTCTGTCTGG 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
LOCUS      R33537      344 bp      mRNA      EST      28-APR-1995
DEFINITION yh79d03.r1 Soares Placenta NB2HP Homo sapiens CDNA clone
            IMAGE:135941 5' similar to SP:VA93_VACC P21062 PROTEIN ;, mRNA
            sequence.
ACCESSION  R33537
NID        9789395
VERSION    R33537.1 GI:789395
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens.
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homiidae; Homo.
            1 (bases 1 to 344)
            Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
            The WashU-Merck EST Project
            Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1078
High quality sequence stops: 275
Source: IMAGE Consortium, LNL.
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone; similarity on wrong strand
Insert length: 1078 Std Error: 0.00
Seq primer: M13Rpl
High quality sequence stop: 275.
Location/Qualifiers
1..344
/organism="Homo sapiens"
/notes="Organ: placenta; Vector: pRT3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st"
```

strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGAGAAATTCGGCGCGAGATTTTATTTTATTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV73 vector. Library constructed by Bento Soares and M. Fatima Bonaldo. "

/db_xref="GDB:541932"
/db_xref="taxon:9606"

/map="11"

/clone="IMAGE:135941"
/clone_lib="Soares placenta NB2HP"

/sex="Female"

/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"

BASE COUNT 74 a 87 c 106 g 75 t 2 others
ORIGIN

Query Match

5.8%; Score 41; DB 30; Length 344;

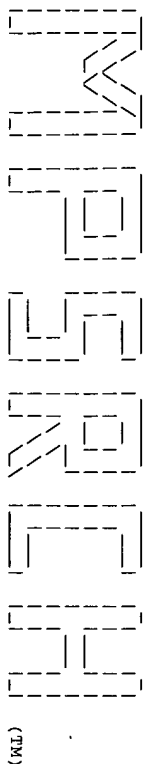
Best Local Similarity 92.3%; Pred. No. 7.18e-34;

Matches 48; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Db 293 TGAACCTTCCCGACCCCTGTCNCCTTGACACATTGCGCACACGGACAC 344

Cp 711 TGAACCTTCCCGACCCCTGTCNCCTTGACACATTGCGCACACGGACAC 661

Search completed: Wed Aug 4 01:29:47 1999
Job time : 1427 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPearch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Aug 4 00:40:47 1999; Maspar time 1317.19 Seconds

Tabular output not generated. 1498.343 Million cell updates/sec

Title: >US-09-240-410-3
Description: (1-712) from US09240410.seq
Perfect Score: 708
N.A. Sequence: 1 CGCGCTGGCGCGGAGGCGCA.....CGAGGTTGGGAGAGTTTCA 712
Comp: GCGCGAGCGGCGGCTCCCGGT.....GTCGCCACCCCTTCAAGTN

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 646147 segs, 1385953633 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

emb158
1:em_ba1 2:em_ba2 3:em_fun 4:em_htg 5:em_hum1 6:em_hum2
7:em_in 8:em_com 9:em_or 10:em_ov 11:em_pat 12:em_ph
13:em_pl 14:em_ro 15:em_sts 16:em_vl
genbank111
17:gb_ba1 18:gb_ba2 19:gb_htg1 20:gb_htg2 21:gb_in1
22:gb_in2 23:gb_com 24:gb_ov 25:gb_pat 26:gb_ph 27:gb_pl1
28:gb_pl2 29:gb_pr1 30:gb_pr2 31:gb_pr3 32:gb_ro
33:gb_st 34:gb_sts 35:gb_sy 36:gb_un 37:gb_vl

Statistics: Mean 10.362; Variance 4.482; scale 2.312

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	620	87.6	2594	AF071542	Homo sapiens semaphori	0.00e+00
2	620	87.6	2636	AF030698	Homo sapiens semaphori	0.00e+00
3	620	87.6	2661	AF068493	Homo sapiens GPR-anch	0.00e+00
4	431	60.9	1192	AF030699	Mus musculus semaphori	0.00e+00
5	431	60.9	2608	AB017532	Mus musculus mRNA for	0.00e+00
6	143	20.2	8861	AF030697	Homo sapiens semaphori	1.14e-102
7	143	6.5	11874	AH018243	Alcelaphine herpesviri	1.60e-16
8	46	6.5	130608	AF005370	Alcelaphine herpesviri	1.26e-09
9	37	5.2	215	128278	Sequence 5 from patent	6.82e-09
10	36	5.1	7218	166494	Sequence 14 from patent	3.64e-08
11	35	4.9	1056	EV04076	Mustela vison GT dinuc	1.91e-07
12	34	4.8	565	EV04076	gDNA encoding envelope	1.91e-07
13	34	4.8	965	AR024229	Sequence 22 from patent	1.91e-07

14	33	4.7	74371	31	AC005369	Homo sapiens chromosom	9.84e-07
15	32	4.5	965	25	AR024229	Sequence 22 from patent	4.96e-06
16	31	4.4	1056	23	EV04076	Mustela vison GT dinuc	2.45e-05
17	31	4.4	216021	31	HUAC004787	Homo sapiens Chromosom	2.45e-05
18	30	4.2	2879	24	AF124485	Danio rerio secreted c	1.18e-04
19	29	4.1	215	25	128278	Sequence 5 from patent	5.57e-04
20	29	4.1	74371	31	AC005369	Homo sapiens Chromosom	5.57e-04
21	29	4.1	216021	31	HUAC004787	Homo sapiens Chromosom	5.57e-04
22	27	3.8	2770	32	AF030430	Mus musculus semaphori	1.14e-02
23	26	3.7	3503	32	S79463	M-sem F-a factor in n	4.94e-02
24	25	3.5	267	17	SL16SRN1	S. lividus gene for 16S	2.08e-01
25	25	3.5	2715	24	GGU28240	Gallus gallus collapsi	2.08e-01
26	25	3.5	10772	21	AF012089	Drosophila melanogaste	8.42e-01
27	24	3.4	66	25	141364	Sequence 144 from patent	8.42e-01
28	24	3.4	69	25	141362	Sequence 142 from patent	8.42e-01
29	24	3.4	2872	32	MMRNS5MA	M.musculus mRNA for se	8.42e-01
30	24	3.4	20521	30	HSMTCHN2	H.sapiens DNA for enha	8.42e-01
31	24	3.4	39540	31	AC000079	Homo sapiens Chromosom	8.42e-01
32	23	3.2	177805	20	AC006450	Homo sapiens Chromosom	3.29e+00
33	23	3.2	69	25	141362	Sequence 142 from patent	3.29e+00
34	23	3.2	565	25	EV04076	gDNA encoding envelope	3.29e+00
35	23	3.2	1719	23	OCHPAMV03	O.cuniculus (OCCDN41)	3.29e+00
36	23	3.2	2993	24	AF073289	Danio rerio transmembr	3.29e+00
37	23	3.2	4113	18	AF028736	Serratia marcescens si	3.29e+00
38	23	3.2	4623	32	CRUKLCG1	Citricellus griseus (Ch	3.29e+00
39	23	3.2	117919	29	HS1184F4	Human DNA sequence ***	3.29e+00
40	23	3.2	132470	29	HS203P18	Human DNA sequence fro	3.29e+00
41	23	3.2	132645	19	AC002345	Genomic sequence from	3.29e+00
42	23	3.2	136124	29	HS394P21	Homo sapiens DNA sequ	3.29e+00
43	23	3.2	141878	30	AC000004	Genomic sequence from	3.29e+00
44	23	3.2	154303	20	AC007189	Homo sapiens, clone 19	3.29e+00
45	23	3.2	174898	19	HS885L7	Human DNA sequence ***	3.29e+00

ALIGNMENTS

RESULT 1
LOCUS AF071542 2594 bp mRNA PRI 01-DEC-1993
DEFINITION Homo sapiens semaphorin K1 mRNA, complete cds.
ACCESSION AF071542
VERSION 93930578
KEYWORDS AF071542.1 GI:3930578

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

FEATURES

CDS

CDS

```

/note="sema K1: glycosylphosphatidylinositol-linked
protein; similar to Alcelaphine herpesvirus 1 semaphorin"
/codon_start=1
/product="semaphorin K1"
/protein_id="AAC80456.1"
/db_xref="PID:93930579"
/db_xref="GI:3930579"
/translation="MTPEPGRRAAPSAREVPGPAPRTGLPIRLILLMAAAS
OGHRSRGPRTFAVKKGYGDDRVDPGAEPEPTVLPHEHGSSSWYGGCKXYLDPEP
GKNAVSRTVNI GSTKGSLDRDCENTITLERREGGLACTGNRHPSCNNLVNGTV
VPLEGRGAPSPDPENSLVLEGEVSTLRKQDNIIFRFRITREISLELYSDY
MNOQFIKATIVHODQADDKIRFEYFRDNDDKDEALNRSVAOLRGDGGESSIS
SVSWKNTLEKMLAVLCSDAATNRKNFNRLDVLPLPPSQMDRTFYGVFSWPNVSA
CVYSGDIDKYFRSTSLKGYHSLLENPPGKLPQOQIPIETFOARHVEVAKORVA
PMGLPKPTLPFSKHYHOKVAVHRMQASGFEFNHYLYTTDGGTIHKYVPEQDESEF
FNMEIIOFPERRAAAIOTMSLDAERKLIVSSQWMSOYPLDCEYKVGCGGCLMSR
PYCGDQGRCLISTISSEKSVLOSINPAPRHECPYKPRDKAPLOKVASLAPSRKLYL
PMESHNATYVSWRHEENVEQSCPEQHSQNLILFENLTAOQYGHFCEADQSGYFRE
QHMDLLPEDGIMAEHLIENHACALASLWGLVLPPTLLGLVH"

```

Query Match	Similarity	96.6%	Score 620	DB 31	Length 2594
Best Local	81	Conservative	0	Pred. No. 0.00e+00	Mismatches 16
Matches	681	Indels	8	Gaps	8
Db 128	CCGCCAGGGCCACCTTAAGAGCGGAGCCCGCATCTTCGCGCTGTGGAAAGGCCATGTAG	187			
Qy 8	CCGCCAGGGCCACCTTAAGAGCGGAGCTTCTGCGCGCTGTGGAAAGGCCATGTAG	67			
Db 188	GGCAGAGCCGGGTGACCTTTGGCCAGACTGAGCCCGCACAGCGGTGCTTTTCCACGAGCCAG	247			
Qy 68	GGCAGAGCCGGGTGACCTTTGGCCAGACTGAGCCCGCACAGCGGTGCTTTTCCACGAGCCAG	127			
Db 248	GCAGCTCTCTGTGTGGGTGGGAGGACGCTGGCAAGGTCTACTCTTTGACTTCCCGGAG	307			
Qy 128	GCAGCTCTCTGTGTGGGTGGGAGGACGCTGGCAAGGTCTACTCTTTGACTTCCCGGAG	187			
Db 308	GCAACACGCACTCTGTGCGCAGCGTGAATATGGCTCCCAAGAGGGTCTGTGTGGATA	367			
Qy 188	GCAACACGCACTCTGTGCGCAGCGTGAATATGGCTCCCAAGAGGGTCTGTGTGGATA	247			
Db 368	AGCGGACTGCGAGACTACATCACTCTCTGGAAGGCGGAGTGAAGGGCTGTGGCCT	427			
Qy 248	AGCGGACTGCGAGACTACATCACTCTCTGGAAGGCGGAGTGAAGGGCTGTGGCCT	307			
Db 428	GTGGACCAACGCGCCCGGACCCCGACCTGCTGGAACCTGTGTAATGACATGTGTGCGAC	487			
Qy 308	GTGGACCAACGCGCCCGGACCCCGACCTGCTGGAACCTGTGTAATGACATGTGTGCGAC	366			
Db 488	TTGGCGAGATGAGAGGCTACGCGCCCTTCACCGCGGAGGAAC-TGCC-TGGTTCTGT	544			
Qy 367	CTTGGCGAGATGAGAGGCTACGCGCCCTTCACCGCGGAGGAACGTCCTGGTTCTGT	426			
Db 545	TT-GAAGGGGAGAGGTGTATTCACCATCCGGA-GCA-GGAAT-ACAATGGGAAGATC	600			
Qy 427	TTTGAAGGGGAGAGGTGTATTCACCATCCGGAAGCAAGAAATTACATTTGGGAAGAT	486			
Db 601	CCTGGTTCCGCGCATCCGGGGCGAGAGTGAAGCTGTATACACAGTATCTTCATGCG	660			
Qy 487	CCTGGTTCCGCGCATCCGGGGCGAGAGTGAAGCTGTATACACAGTATCTTCATGCG	546			
Db 661	AACCCACAGTTTCATAAACCCACATCGTGACCAAGACACAGGCTTACGATACAGATC	720			
Qy 547	AACCCACAGTTTCATAAACCCACATCGTGACCAAGACACAGGCTTACGATACAGATC	606			
Db 721	TACTACTTCTTCCGAGAGGACATCCTGTCAAGAAATCCTGAGGCTCTTCATGTGTCC	780			
Qy 607	TACTACTTCTTCCGAGAGGACATCCTGTCAAGAAATCCTGAGGCTCTTCATGTGTCC	666			
Db 781	CGTGTGCGCCAGTTGTGCAGGGGGGACCAAGGTTGGGAAAGTTCA 825				
Qy 667	CGTGTGCGCCAGTTGTGCAGGGGGGACCAAGGTTGGGAAAGTTCA 711				

[illegible]

D	b	320	GCAAAAGCATCTGTGGCAGCGGAATATGCGCTCCACAAAGGGGTCCTTGGAATA	379
O	y	188	GCAAAAGCATCTGTGGCAGCGGAATATGCGCTCCACAAAGGGGTCCTTGGAATA	247
D	b	380	AGCGGACTGCAGAACTACATCACTCTCTTGGAGAGCGGAGTAGAGGGTCGTGGCCT	439
O	y	248	AGCGGACTGCAGAACTACATCACTCTCTTGGAGAGCGGAGTAGAGGGTCGTGGCCT	307
D	b	440	GTGGCACCAAGCCCCGGCACCCCAGCTGCTGGAACCTGTTGATATGGACTGTGTGCCAC	499
O	y	308	GTGGCACCAAGCCCCGGCACCCCAGCTGCTGGAACCTGTTGATATGACTGTGTGCCAC	366
D	b	500	-TTGGCGAAGTAAAGGCGTAAGCGCCCTTCACCCCGGAGAAGAAC-TCCC-TGGTTCGT	556
O	y	367	CTTGGCGAAGTAAAGGCGTAAGCGCCCTTCACCCCGGAGAAGAACCTCCGTGTGTCTGT	426
D	b	557	TT-GAAGGGAGCAGAGGTGTATTCCACCATCCGGA-ACAAT-ACAATGGGAAGATC	612
O	y	427	TTTGAAGGGAGCAGAGGTGTATTCCACCATCCGGAAGCAAGGAATTACATTGGGAAGAT	486
D	b	613	CCTCGGTTCCGCCGATCCGGGGCGAAGTAGCTGTACACAGTATCTGTATGCAG	672
O	y	487	CCTCGGTTCCGCCGATCCGGGGCGAAGTAGCTGTATACACAGTATCTGTATGCAG	546
D	b	673	AACCCAGATTATCAAAGCCACACATCGTGGCACCAGACACAGGCTTACATACAAAGATC	732
O	y	547	AACCCAGATTATCAAAGCCACACATCGTGGCACCAGACAGGCTTACATACAAAGATC	606
D	b	733	TACTACTTCTTCCGAGAGGACATCCTGTACAAGAAATCCTGAGGCTCTCTCAATGTGCC	792
O	y	607	TACTACTTCTTCCGAGAGGACATCCTGTACAAGAAATCCTGAGGCTCTCTCAATGTGCC	666
D	b	793	CGTGTGCCCCAGTTGTGCAGAGGGGGACACAGGGTGGGAAAATTCA	837
O	y	667	CGTGTGCCCCAGTTGTGCAGAGGGGGACACAGGGTGGGAAAATTCA	711

RESULT 3

LOCUS	AF069493	2661 bp	mRNA	PRI	03-DEC-1998
DEFINITION	Homo sapiens GPI-anchored membrane protein CDw108 precursor, mRNA,				
ACCESSION	AF069493		complete cds.		
NID	g3947939				
VERSION	AF069493.1	GI:3947939			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
AUTHORS	Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 2661)				
JOURNAL	Yanada,A., Kubo,K., Takeshita,T., Haraashima,N., Kawano,K.,				
FEATURES	Sagawa,K., Sugamuru,K. and Itoh,K.				
source	Direct Submission				
	Submitted (02-JUN-1998) Immunology, Kurume University School of Medicine, Asahi-machi 67, Kurume, Fukuoka 830, Japan				
	Location/Qualifiers				
	1..2661				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="15"				
	/map="between WI-6247 and CHLC.GATA85D02"				
	/cell_line="HPB-ALL"				
	/cell_type="T lymphocyte"				
	18..2018				
	/note="putative adhesion molecule; lymphocyte activation antigen"				
	/codon_start=1				
	/product="GPI-anchored membrane protein CDw108 precursor"				
	/protein_id="AAC82642.1"				
	/db_xref="PID:g3947940"				
	/db_xref="GI:3947940"				
	/translation="MTPTTPGRRAAPSARNAVPGPPARLGLPLRLILLIMAAAAASA"				

	Query Match	Best Local Similarity	Matches	87.6%;	Score 620;	DB 31;	Length 2661;
Db	145	CCGCCACGAGGCCACCTAAGAGAGCGGAGCCCGACCTTCGCCGCTGTGGAAGGCCATAG	204				
Qy	8	CCGCCACGAGGCCACCTAAGAGAGCGGAGCCCGACCTTCGCCGCTGTGGAAGGCCATAG	67				
Db	205	GGCAGAGACCGGGGTGACTTGGCCAGACTAGCCGACAGCGTGCCTTCCACAGCAG	264				
Qy	68	GGCAGAGACCGGGGTGACTTGGCCAGACTAGCCGACAGCGTGCCTTCCACAGCAG	127				
Db	265	GCACCTCCTCTGTGTGGGTGGAGAGAGTGGCAGAGCTACTCTTGTACTTCCCGCAGG	324				
Qy	128	GCACCTCCTCTGTGTGGGTGGAGAGAGTGGCAGAGCTACTCTTGTACTTCCCGCAGG	187				
Db	325	GCAGAGACGCATCTGTGTGGCAGCGTGAATATCGGCTCCACAAAGGGTCTGTGTGATA	384				
Qy	188	GCAGAGAGCGCATCTGTGTGGCAGCGTGAATATCGGCTCCACAAAGGGTCTGTGTGATA	247				
Db	385	AGCGGAGACTGGGAATCACTACATCTCTCCGAGAGAGCGGAGTGAAGGGCTGTGGCCT	444				
Qy	248	AGCGGAGACTGGGAATCACTACATCTCTCCGAGAGAGCGGAGTGAAGGGCTGTGGCCT	307				
Db	445	GTGGCACCAACGCCCGGACCCCGAGCTGTGAACCTGTGTAATGACATGTGGTGCAC	504				
Qy	308	GTGGCACCAACGCCCGGAGCCCGAGCTGTGAACCTGTGTAATGACATGTGGTGCAC	366				
Db	505	-TTGGCGAGATGAGAGGCTACGCCCTTACAGCCCGAGAGAAC-TCCC-TGGTTCTGT	561				
Qy	367	CTTGGCGAGATGAGAGGCTACGCCCTTACAGCCCGAGAGAACGTCCGTGGTTCTGT	426				
Db	562	TT-GAAGGGAGCGAGGTGTATTCACATCCGGA--GCA--GGAT--ACATATGGGAATATC	617				
Qy	427	TTTGAAGGGAGCGAGGTGTATTCACATCCGGAACCAAGGAATTTCAATTTGGGAAGAT	486				

Query	Match	60.9%	Score 431	DB 32	Length 1192	
Db	618	CTCGGTTCCGCCGATCCGGGGGAGAGAGTGTGTACACCACTGATCTACTGTATCAG	677			
Db	487	CTCTGGTTCCGCCGATCCGGGGGAGAGTGTGTACACCACTGATCTACTGTATCAG	546			
Db	678	ATCCCACTGTTCAATCAAGGACCATGCTGTCACCAAGACAGGCTTCATGTACAGATC	737			
Db	547	ATCCCACTGTTCAATCAAGGACCATGCTGTCACCAAGACAGGCTTCATGTACAGATC	606			
Db	738	TACTACTTCTTCGAGAGGACCAATCTCTGACAGATCTTGAAGCTCTCTCAATGTGTC	797			
Db	607	TACTACTTCTTCGAGAGGACCAATCTCTGACAGATCTTGAAGCTCTCTCAATGTGTC	666			
Db	798	CGTGTGACCCAGTTGTGTACAGAGGAGGACAGGAGTGGGAAAGTTCA	842			
Db	667	CGTGTGACCCAGTTGTGTACAGAGGAGGACAGGAGTGGGAAAGTTCA	711			
RESULT	4	AF030699	1192 bp	MRNA	ROD	04-SEP-1998
DEFINITION		Mus musculus semaphorin 1 (Semal)		MRNA, partial	cds.	
ACCESSION		AF030699				
MD		93523116				
VERSION		AF030699.1	GI:3523116			
KEYWORDS						
SOURCE						
ORGANISM		house mouse.				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS		1 (bases 1 to 1192)				
TITLE		Lange, C., Liehr, T., Goen, M., Gebhart, E., Fleckenstein, B. and Ensser, A.				
JOURNAL		New eukaryotic semaphorins with close homology to semaphorins of DNA viruses				
MEDLINE		Genomics 31 (3), 340-350 (1998)				
REFERENCE		98389619				
AUTHORS		2 (bases 1 to 1192)				
TITLE		Ensser, A., Liehr, T., Lange, C., Goen, M., Gebhart, E. and Fleckenstein, B.				
JOURNAL		Direct Submission				
FEATURES		Submitted (21-OCT-1997) Institut fuer Klinische und Molekulare Virologie, Friedrich Alexander Universitaet, Schlossgarten 4, Erlangen, Bavaria 91054, Germany				
source		Location/Qualifiers				
gene		1..1192				
CDS		/organism="Mus musculus"				
		/db_xref="taxon:10090"				
		/map="9A.3-B"				
		/chromosome="9"				
		1..>1192				
		/gene="Semal"				
		/evidence=not_experimental				
		13..>1192				
		/gene="Semal"				
		/note="M-Sema-L"				
		/codon_start=1				
		/evidence=not_experimental				
		/product="semaphorin 1"				
		/protein_id="AAC34262.1"				
		/db_xref="PID:93523117"				
		/db_xref="GI:3523117"				
		/translation="MTPPPGRAPASAPARAVLSLPARFGLPLRLILVPMVAASAS				
		OGHSISGRISAVMKGODHPDOEPHTVLFHEBGRSVAVGGKGYHNPFGKRS				
		ASRVIVNIGSTKSGODKODGNYITLLERNGILVCGTAAKRPSCNLYVNDYSVMS				
		LGEMKGYAPSPDENSILVLFEGDEYITIRKQETNGKIPRRRIIGSEELITSDTVWC				
		NPOFKATIVHODAVIDKIIYFFREDNPKNPEAPLVNPGSVVAOLCRGDGESSLS				
		SKMWFLEKAMLVCSDAATNRNFRLODFELLPDPSGQWRDTRVYGSNPNVYAVCC				
		YSLGIDFIVERTSSLYKGYHMLNSPRPCKPKPKPIPTFTFYVADSHPEAVQAVEPR				
		G"				
BASE COUNT		254 a	355 c	336 g	247 t	
ORIGIN						

Match	562: Conservative	0: Mismatches	77: Indels	9: Gaps	9:
Best Loca Similarity 86.7%: Pred No. 0.00e+00:					
Db	189	AGGGCAGGACCATGTGGACTTTAGCCAGCCTGAGCCACACACCCTGCTTTTCATGAGCC	248		
Qy	66	AGGGCAGGACCGGGGTGACTTTGGCCAGATGAGCCGACACAGGCTGTTTCCACGAGCC	125		
Db	249	GGGCACTTCTCTGTGTGGGTGGAGCTGGCAGAGTCTACCACTTCAACTTCCCCCA	308		
Qy	126	AGGCACTCTCTGTGTGGGTGGAGAGCTGGCAAGGTCTTACTCTTTGACTTCCCCG	185		
Db	309	GGGCAAGATGCTCTGTGGCAGCGGTGAACATCGGCTCCACAAAGGGGCTGTGAGGA	368		
Qy	186	GGGCAAGAAAGCATCTGTGGCAGCGGTGAATATCGGCTCCACAAAGGGGCTGTGGA	245		
Db	369	CAAAACGAGCTGTGGGAATTACATCACTCTTCTAGAAAGCGGGGTAAATGGGCTGTGT	428		
Qy	246	TAGCGGCACTGGGAGAACTACATCACTCTCTGAGAGCGGAGTGAAGGGCTGTGCG	305		
Db	429	CTGTGGCACCACATGGCCCGGAGGCCAGCTGCTGGAATCTTGGAATGACAGTGGTGAT	488		
Qy	306	CTGTGGCACCACAGCCCGGACCCCACTGTGGAACCTGTGGAATG-CACGTGGTGGCC	364		
Db	489	GTCACCTTGGAGATGAAGAGCTATCCCTCCCTCACCCGGATGAAAC-TCGC-TGGGT	546		
Qy	365	A-C-CTTGGCAGAGTGGAGGCTACGCCCTTTCACCCCGGACGAGAAAGTCCCGTGT	422		
Db	547	CTGTGT-GAAGAGATGAAGTACTTACTCTACATCCGGA-AGA-GGAAT-ACAAAGGGA	602		
Qy	423	CTGTGTGAAGGGGAGAGAGTGTATCCACCATCCGGAAGCAGGAATTTACATTTGGA	482		
Db	603	GATCCCTCGGTTTCGACGCAATTCGGGGCAGAGTGAATCTGTACACAAGTGAATCACTCAT	662		
Qy	483	AGATCTCTCGTTCCGGCCGATCCGGGGCAGAGTGAAGTGAAGTACACAGTGAATCTCAT	542		
Db	663	GCAGAACCCACAGTTTCATCAAGGCCACATTTGACCAACCAAGCCTATGATGATAA	722		
Qy	543	GCAGAACCCACAGTTTCATCAAGGCCACATCTGTGACCAACCAAGCCTATGATGATCA	602		
Db	723	GATCTACTACTTCTTCCGAGAGACCAACCTTACAAGAACCCCGAGGCTCTCTCAATG	782		
Qy	603	GATCTACTACTTCTTCCGAGAGACCAATCTTACAAGAAATCTGTAGAGCTCTCTCATATG	662		
Db	783	GTCCCGAGTAGCCCAAGTTGTGCAGGGGGGAGCAGAGGTGTGAGAGTTC	830		
Qy	663	GTCCCGGTGTGGCCAGTTGTGCAGGGGGGAGCAGAGGTGTGAGAAAGTTC	710		
RESULT	5				
LOCUS	AB017532	2608 bp	mRNA		
DEFINITION	Mus musculus mRNA for msemk1p, complete cds.				
ACCESSION	AB017532				
NID	94519587				
VERSION	AB017532.1	GI:4519587			
KEYWORDS	msemk1p.				
SOURCE	Mus musculus postnatal brain cDNA to mRNA.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
AUTHORS	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (sites) and Takahashi, H.				
	Sato, Y. and Takahashi, H.				
	Molecular cloning and expression of novel semaphorin family of				
	molecule				
JOURNAL	unpublished (1998)				
REFERENCE	2 (bases 1 to 2608)				
AUTHORS	Takahashi, H. and Sato, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (08-SEP-1998) to the DDBJ/EMBL/Genbank databases.				
	Takahashi, Mitsubishi Kasel Institute of Life Sciences,				
	Neurobiology Lab; 11 Minamiloopa, Machida, Tokyo 194, Japan				
	(E-mail: hiroshi@elabra.is.m.kagaku.co.jp, tel: 81-427-24-6211,				
	Fax: 81-427-24-6314)				
FEATURES	Location/Qualifiers				

	Db	2176	ACCCAGATTCAATCAAGCACCACATGCTGGACCAAGCACGGCTTAGATGACAACATCT	2235	
	Oy	548		607	
	Db	2236	ACTACTTCTTCGAGAGGACAATCTGCAGAAGATCTCGAGCTCCCTCATAGTGCC	2295	
	Oy	608		667	
	Db	2296	GTTGTGGCCCCAGTTTGTCTCAGGTTGAACACGGG	2326	
	Oy	668		698	
			GTTGTGGCCCGATTGTGTCAGAGGGGACCAAGG		
RESULT LOCUS DEFINITION ACCESSION NID VERSION KEYWORDS SOURCE ORGANISM REFERENCE TITLE JOURNAL MEDLINE REFERENCE AUTHORS TITLE JOURNAL FEATURES source repeat_region cds	7	AH018243	11874 bp DNA	VRL 30-SEP-1995 Alcelaphine herpesvirus 1 putative semaphorin homolog (AHV-sema) and putative membrane antigen genes, complete cds, and major ssDNA-binding protein gene, partial cds. U18243 g10000716 U18243.1 GI:1000716	
			Alcelaphine herpesvirus 1.		
			Alcelaphine herpesvirus 1.		
			Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; unclassified Gammaherpesvirinae.		
			1 (bases 1 to 11874)		
			Enser,A. and Fleckenstein,B.		
			Alcelaphine herpesvirus type 1 has a semaphorin-like gene		
			J. Gen. Virol. 76 (Pt 4), 1063-1067 (1995)		
			2 (bases 1 to 11874)		
			Enser,A.		
			Direct Submission		
			Submitted (07-DEC-1994) Armin Enser, Institut fuer Virologie,		
			Friedrich-Alexander-Universitaet Erlangen-Nuernberg, Schlossgarten		
			4, Nuernberg, Bavaria 91054, Germany		
			Location/Qualifiers		
			1..11874		
			/organism="Alcelaphine herpesvirus 1"		
			/strain="C50"		
			/specific_host="Connochaetes taurinus taurinus"		
			/db_xref="taxon:35252"		
			/lab_host="Bos taurus"		
			/clone="clones C238, C249, C500BS"		
			<1..(330.346)		
			/note="gamma-2 herpesviral GC-rich terminal tandem repeats"		
			/function="putative packaging signals"		
			/rpt_type=tandem		
			3573..5534		
			/standard_name="AHV-sema"		
			/note="putative semaphorin homolog; similar to Genbank Accession Number U26081"		
			/codon_start=1		
			/protein_id="AAC54475.1"		
			/db_xref="PID:g1000717"		
			/db_xref="GI:1000717"		
			/translation="MAYLNATAIVKSPVISLSLSTKKVLKEFGSGEGQGLGLEFVIH		
			PAAAGTLCVSTRILMLISATIAKSRFLDKRLLVNLTDGGRFRGPDPHYLVFR		
			SLNSDYYVGNNNTIYLFDFAHSSNASTALINISTHTRLSSTCEAFITLINQED		
			GLGATNSOKPSQCMWLNINLTOTFLGKLGAAPSSGNLVLPDDNDYTILNKYS		
			LSGHKRRRIAGIAGELYETSDAMRPQFOVAIAVKHNSYDKTYFFEOSSHDFQR		
			EPRHVPRVGOVCSDDOGSESTSYKKRTPELKALADVDYDTGAIYNELDIPTQAR		
			ENSMERELLIGLEPESPMNFSAVCFTYTKKDIDHVEIKSKLNYHHKLTPRGQQCKNH		
			OHVFTEFOVADRIPEVADVYKNNAEPFIQSRYITKLAVRVGEVFAATIFY		
			LTTLKGITHIVAREDENSTITALNILEINPQKPATIONILLDNINKLYNSEMS		
			EVPLDLCSYVNGDFSCFMRSDELCTWNNGCSFKORSYVTGGPARNTISEMGDAH		
			APTVAKHOVSPIPLSNSYLSCPASNADYFWMDGFTEKRCHKTAKNDICILLANS		
			TTLANGHCNMKRDSTYVKILEVNTLM"		
			6414..10523		
			/note="putative membrane antigen; similar to p140 (PIR		
CDS					

[illegible]

Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae.

REFERENCE 1 (bases 1 to 130608)

AUTHORS Essner, A., Pflanz, R. and Fleckenstein, B.

TITLE Primary structure of the alcelaphine herpesvirus 1' genome

JOURNAL J. Virol. 71 (9), 6517-6525 (1997)

MEDLINE 97404659

REFERENCE 2 (bases 1 to 130608)

AUTHORS Essner, A., Pflanz, R. and Fleckenstein, B.

TITLE Direct Submission

JOURNAL Submitted (27-MAY-1997) Institut fuer Klinische und Molekulare Virologie, Friedrich-Alexander-Universitaet Erlangen-Nuernberg, Schlossgarten 4, Erlangen 91054, Germany

location/Qualifiers

source 1. 130608

/organism="Alcelaphine herpesvirus 1"

/strain="C500"

/specific_host="Connchaetes taurinus taurinus"

/db_xref="taxon:35252"

/lab_host="MDBK bovine epithelial kidney cells; ATCC CCL-22"

918. .1211

/note="putative ORF"

/codon_start=1

/evidence=not_experimental

/product="A1"

/protein_id="AAC58052.1"

/db_xref="PID:g2337968"

/db_xref="GI:2337968"

/translation="MRLHFLFYRSIEFYMKSKHKIFCITLSLYFLTAHKRRARPITDFSRIMGCPCEFKMSRHOLLNLQMLSPCRKCSLTSADOCITANEDITK"

1248. .1254

/evidence=not_experimental

complement(1300..1305)

/note="spliced cDNA detected by 5' RACE-PCR"

/codon_start=1

/evidence=experimental

/product="A2"

/protein_id="AAC58053.1"

/db_xref="PID:g2337969"

/db_xref="GI:2337969"

/translation="MSQNSNSNPSPKRRKYKMKDLETOEKERRSINRASKNFKRRRIEEOEKGLINKENSRILRCOVERKRDRLREWLNYHKCTTLQNYNGPPEPRKVENSLKMCATATFLNDQYTTNNLIPVSGSNTNGFAATATLHNCYCKTLANNINFEAKLNCCEVLPSTSLDLDLSIDWNNTLNL"

3492. .5453

/note="A3; AHV-sema, similar to Vaccinia A39"

/codon_start=1

/evidence=not_experimental

/product="semaforin homolog"

/protein_id="AAC58054.1"

/db_xref="PID:g2337970"

/db_xref="GI:2337970"

/translation="MAYLNATVSPVLSLSKVLKFEHGGEGGGLITFEVTHPAMGTCVSRILMLISATTAARSPIDKPRILVNTLDCGGRHSGRPOEPRLEHSLNSDYYVCGNNTITLFDFAHSSNASTALININTSTNTRLSCTGCTTLHNOJDGLLACGTNSQKPSGCWLNINLTTOFLGKGLDAPSPSSGNLVLEDDONTSTINLXSLGSHKRRIRAGVDELYTSDTAMHRPOVQATVHKNSYDCKIYEPFONSNDYKSPHTVPRVGVCSDDOGESSLSVYKWTFLKARLACVDYDTRIKYELQDIFIMQAPENSMEEETLIYGLFSPNFSFACVETKIDHVEKTSKLKNHKKLPTPRGCGMKRHQHVTEFQVADRYPRVADPVQKKNAMPRIIOSKYTKLYIRVRYGCVFPAATIFYLTTIKGTIHLIYRREDSNTALILEINFQKRAPQIIONLLDNTNKLKLVNSEWYS EVPDLDCSVGNDCSCFEMSRLPCTWYNTGSCFQKQSVETGPGPARILTSEMGDHYAPTGVKHQVSIPLISNYSLSCPAVSNADYFWTKDGETEKRCVYKHKNDICILLIANS TTATNGTHVCNMKEDSVTKLLEVNITLM"

5590. .5597

5732. .6097

/note="ORF containing putative signal peptide; detected by 5' RACE-PCR"

/codon_start=1

/evidence=experimental

/product="A4"

polysignal

/protein_id="AAC58055.1"

/db_xref="PID:g2337971"

/db_xref="GI:2337971"

/translation="MVAOLYHNLITLIASIIYIFFNAPATLYEDDEDFNSYEVFLSDSRQFPRRNSCNCCTREATRTILLEALKRALQIIAGIINETDIELPPTPIWTMTPLTTPLDTSPPVLPSPAIp"

6123. .6128

/evidence=not_experimental

6333. .10442

/note="ORF03; similar to H. saimiri and EHV2 ORF3, similar to ORF75"

/codon_start=1

/evidence=not_experimental

/product="tegument protein"

/protein_id="AAC58056.1"

/db_xref="PID:g2337972"

/db_xref="GI:2337972"

/translation="MTRRELQLAVTAEPSSALETAIAALRSIPGLONVITOTADVLVTFPTSPRPRLRHQKLETLFLFKALRTFDEQOYPIPLSALSSFTFYVGPIDNRLPTQSNEIATLIOTRDAREKVEVVERIHCRLFFSGSGEOLNHTHYALREILCGDLTXOQPRNLEPNSLDFATILPPREILDKVRYHYVHGDINLTGAASVYNASPHATYDSEIFRPIPTHVYRFDIMEIRVAVPDVGHSGRVFRHOGCAHYCSTLNSTGATGCLQKGNITISISQPTLGQLITTEPCGCGQQLSATATQMLTKTANVETAKTINYAGVPLVOGFVLTSPTLREKRAAFGSLITGLPIYLIPTAETLDAMERNRQLLVAGYEDYEGKRLHNPVAINMNSSESRHAHILKQALROLMIIPFGVHICCDM QYNEKVTMEKIAKACGDLGMSARSILPHTTLRLKMNERNLNYHRIIKNSMLKVD SAAIVVIAIDNEYODYSKKMIDALSGMCGPGLITGLTPSNITVYISDKNOGEIYDI QKMHMPKOPSEGETDPTLAPDMSNTQLRKNDVTDLDLLOVLRHPTVCKAHITAHV DRCNGHIAQOPGVGPDIPLCDFSVYHNLDVGDIREGESVPRVAVADWVARRLI EITYSTPGDITDATALANNGLYHLPSEHQVEKNGCIGIEKTFETORDPL GTILAVSECTNCLIGPVENYEFLGLSISVPEGIHYREIVNSIIMAKKDFCSKNF GFOVNSAENGNCILRSVATAANAPCVPPGSLPYPKGSAITLRYNLHREHLSGICMAASGISAETPTPTPSOLRNLOEMLVYKNAENLISGDSVADGGLICAVCMFAG GLSARLIHDEDEEYFPPLPSETPGVLEYNADVAITARANLYVECOIEYVES DITVYHQNTQLSVSPVSRKHNWTLFSVDLLVYKEQVLEETSYGIEVHLVAD PYSLIIOSTPRPVLVHLPGCGYDPAALALNNSCFSPDTPVPCCKYRNHREDA PGPSPADDEFIAGIYLVGSNIDSVDGDSITROMLVNNOVYINVRNLAKAGSFTLAI GOLACRLIFATKAIIGFDAGSOQTPFLPNASRYESRWLFEKIPEDTKAAVFRDLAC VLPSCWOGTHLGFSSHNNTPEFGDLEEROOVAATPVOSGAPAREVLPTEAEHPY AGCSDGRLHALPDPCLAFNTWQOHQNOTGCGQSELYSPKMLFRLYNKSKHQ HYSRLRLTNLRHTFNETQQLDPHADRRGAVPDNPQYIPMDAMPDSQ"

11134. .11139

/evidence=not_experimental

11260. .14643

/note="ORF06"

/codon_start=1

/evidence=not_experimental

/product="major ss DNA binding protein"

/protein_id="AAC58057.1"

/db_xref="PID:g2337973"

/db_xref="GI:2337973"

/translation="MALKHQLNHAVEDNIGSKAPIGPCGFIIYPTFHNFKESILLG DKYAEAGAFSLPLHGVTEAEVAPVNVKAVYKIDITVTSKISTYNYRAIIFHNK FESIFSGPGLSGLEKACDLFGVPTPLGEGSDVSDLPCKPWOQKIDMAVITEG FKRLHLGLIKLYKSGMSVMINKTIVRPLTDEDLFTKCSRLRLYIPAVEEYLY TLTYSAGSLRVHNASLVAIOEOCPHVKYKAKIKYSREKVALVAGCDTTLMIAD AYAAEGLSTLSFEPARPOEKTIVQDYISMDTIFASCESTSDREALKSKNADLHIIH AOLFEVNSIYVNRVARQAPITPSKVEPVNYSYLYQHGIANCEETLEDGSPATG APASLSLDSGSEFTLOHLYAAEPNPLARCYLQCGHQKSTLNDAYNITVEVGAAN NSVSCSLGQGCQVCIINTLYRLKRPFPVLOGSRDYPVILGINVINEDELFLEF ASRDEDEONTOEETPRYTNOLNITLEKLEAALVSPVADEGAGSGSNLEKF VRRFSDILVLAEAKPTINTMKNNVNRKESIKGSHVIOVNCNMYMAPCSIMLNTLYRSILITTIODIALPISTYSESNPAGCKRPENMLLHTQTLNTRKSFIDKGVITG TEKKVYHAQPSDFEVDADATNNMSYPVAVOURLAOLVALKNIKVKNRILFSPSC SEHYQVAFKTNARBNXYLAGYVFLNSFHQLEPNKISCLYLMASNCKRQKQJPC VEPVSAEALNKEFSYINNNSKOFEEVYNSVDEVDYVYAKQILNNAIIRACQOTAFY AVTIIHSIFPKVOETCALEYVHYLGESVMSADYVYVANNONLAKALYINSSIRASAAILA RSRPIVTLPVVYKXYGAGNOALFOSALYGFPMGSGVKNLIGLGFYKKOONSMR KRYLEMTPLVGNLKRSTYHOGAIEFIEVYKRIQISIILEDQADEVLNRYVCEIAYSL GAGCADLTDLDIOFIIGSTYGMFSENLELVDLQRELVGPMTHEMASVLSKGCETDE YQVFVAEEEOVKILTSDHSGKVVAGKRRIAITMFDLDL"

14725. .16767

/note="ORF07"

polysignal

CP 367 GGTGCACACAGTCATTCACAGGTTCCAGAGCTGGGGTGGCGGCTGTGGCCAC 308
Db 1244 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1303
CP 307 AGGCACGACGCCCTCTACTCCGCTCTCCAGAGAGTATAGTTCTCCGAGTCCGCT 248
Db 1304 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1363
CP 247 TATCCAGACAGAGCCCTTTGTGGAGCCGATATTCACCGTCGACAGATGCTTCTGC 188
Db 1364 YYY 1367
CP 187 CCTC 184

RESULT 11
LOCUS MWU87256 1056 bp DNA MAM 02-JAN-1999
DEFINITION Mustela vison Gt dinucleotide repeat, chromosome 1q.
ACCESSION U87256
NID 94099442
VERSION U87256.1 GI:4099442
KEYWORDS
SOURCE American mink.
ORGANISM Mustela vison
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Carnivora; Fissipedia; Mustelidae; Mustela.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Brusaard, K., Shukri, N.M., Malchenko, S., Koroleva, I. and Lohi, O.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1997) Breeding and Genetics, Danish Institute of
Animal Science, Blichersalle K25, Tjele 8830, DK
FEATURES
source
1. 1056
/organism="Mustela vison"
/db_xref="taxon:9667"
/chromosome="1"
/map="1q"
/note="primers: 1167F: agccctgcatactactctt, 1167R:
gagatctaccgcgtgttag"
primer_bind
98..119
/standard_name="1167F"
/complement(300..320)
primer_bind
/standard_name="1167R"
BASE COUNT 211 a 221 c 210 g 225 t 189 others
ORIGIN

Query Match 4.9%; Score 35; DB 23; Length 1056;
Best Local Similarity 19.7%; Pred. No. 3.64e-08;
Matches 30; Conservative 66; Mismatches 55; Indels 1; Gaps 1;

Db 559 GMSGCKMTWTSGCCSCMRDVBMMKMKWCKSKCKSTGCKDKMSGCAYGKMRRCR 618
QY 292 GAGGGGCTGCTGGCTGTGGACCAAGCCCGGACCTCTGGAACCTGTAAT 351
Db 619 YMMARMCVSGSTWARRCCCKDKSKGSHCKSRKRYKMDRYBCKSMCAMVGVKWSAG 678
QY 352 GCACTGTGGTGGCCACTTGGGAGAGTGGAGGCTA-CGCCCTTCAGCCCGAGAGAA 410
Db 679 TGCARAGWRWDSCTSTKGARRKGMARRK 710
QY 411 CGTCCGCTGTTCTGTTTGAAGGAGAGCAAG 442

RESULT 12
LOCUS E04076 565 bp RNA PAT 29-SEP-1997
DEFINITION gDNA encoding envelope region of type C hepatitis virus.
ACCESSION E04076
NID 92172286
VERSION E04076.1 GI:2172286
KEYWORDS JP 1992349885-A/1.
SOURCE Hepatitis C virus.
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

REFERENCE 1 (bases 1 to 565)
AUTHORS Moriaga, T., Chayama, K., Kumada, H. and Ichikawa, Y.
TITLE NUCLEIC ACID FRAGMENT WITHIN ENVELOPE REGION OF HEPATITIS C VIRUS
AND METHOD FOR DETECTING THE SAME
JOURNAL Patent: JP 1992349885-A 1 04-DEC-1992;
TEIJIN LTD
COMMENT OS Hepatitis C virus
PN JP 1992349885-A/1
PD 04-DEC-1992
PF 29-MAY-1991 JP 1991152169
PI MORINAGA TSUTAE, CHAYAMA KAZUAKI, KUMADA HIROMITSU, PI
ICHIRAWA YATARO
PC C12N15/10, C12O1/68, C12O1/70//C12N15/11;
CC strandedness: Single;
CC topology: Linear;
FEATURES
source
1. 565
/organism="Hepatitis C virus"
/db_xref="taxon:11103"
BASE COUNT 60 a 93 c 107 g 85 t 220 others
ORIGIN

Query Match 4.8%; Score 34; DB 25; Length 565;
Best Local Similarity 27.5%; Pred. No. 1.91e-07;
Matches 44; Conservative 54; Mismatches 61; Indels 1; Gaps 1;

Db 227 CBGGGTGRTBCTGTGTCGSGARRRYNNYTBMGNTGTGGGTCCTBACYC 286
QY 255 CTGGGAGACATACATCTCTCCGAGAGCGAGAGGGGCTGTGGCTGGCAC 314
Db 287 CCAGSTYGCVCRCYARAVYAVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 346
QY 315 CAACGCCGACCCACCCAGCTGCGAA-CCTGATGATGCACTGTGTGCACCTGGCG 373
Db 347 TCGTGTGGGGRSVGCYRCYTCGTGTCGSCGYVYTAATGT 386
QY 374 AGAGTGAAGGCTAGCCCTCTAGCCCGAGAGCAAGAGT 413

RESULT 13
LOCUS AR024229 965 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 22 from patent US 5795961.
ACCESSION AR024229
NID 93977523
VERSION AR024229.1 GI:3977523
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 965)
AUTHORS Wallace, T., Paul, Harris, W.J., Carr, F.J., Old, L.J., Welt, S. and
Kilamura, K.
TITLE Recombinant human anti-Lewis b antibodies
JOURNAL Patent: US 5795961-A 22 18-AUG-1998;
FEATURES
source
1. 965
/organism="unknown"
BASE COUNT 192 a 170 c 226 g 205 t 172 others
ORIGIN

Query Match 4.8%; Score 34; DB 25; Length 965;
Best Local Similarity 14.9%; Pred. No. 1.91e-07;
Matches 26; Conservative 80; Mismatches 66; Indels 2; Gaps 2;

Db 783 RHTVHSGVYSTCTASPTTSYWG-WYRGWGDYGGGTYNNYNGRVTMTADTSSN 841
QY 226 ACAAGGGCTCTGTCTGATTAAGCGGAGTGGAGACATACATCTCTCCGAGAGG 285
Db 842 SRSSVTRADTAIVYCYRGRASVDSGDGYMGTTVYVSSHVVKMTSSSSASVDRYTCR 901
QY 286 CGAGAGTGAAGGCTGTGGCTGTGG-CACCAAGGCCCGGACCCAGCTGTGTGAACCT 344


```

repeat_region /rpt_family="Alu" complement(25727. .26471)
/rpt_family="Alu"
repeat_region /rpt_family="Alu" 27191..27477
/rpt_family="Alu"
repeat_region /rpt_family="Alu" 27774..28057
/rpt_family="Alu"
repeat_region /rpt_family="Alu" 28040..28066
/notes="(A)27"
/rpt_type=tandem
/rpt_unit=A
28712. .28930
/rpt_family="MER20"
misc_feature complement(28769. .28838)
/notes="GRAIL 2 excellent exon, frame 0"
repeat_region 28987..29214
/rpt_family="Alu"
/rpt_family="Alu"
join(29485. .29599,29598. .29779)
/notes="95% identity dbSTS:G14522 (SHGC-11312)"
29495. .29976
/standard_name="A1027942"
misc_feature /notes="100% identity EST ov84a10.x1"
complement(30401. .30363)
/notes="GRAIL 2 excellent exon, frame 2"
repeat_region complement(30682. .30733)
/rpt_family="MIR"
complement(31573. .31724)
/misc_feature /notes="GRAIL 2 excellent exon, frame 1"
complement(32159. .32352)
/notes="GRAIL 2 excellent exon, frame 2"
33388. .32488
/rpt_family="MLT1"
32617. .32908
/rpt_family="Alu"
32877. .33068
/rpt_family="MLT1"
complement(33670. .33785)
/rpt_family="Alu"
/misc_feature complement(34021. .34144)
/notes="GRAIL 2 excellent exon, frame 2"
complement(35238. .35331)
/notes="GRAIL 2 excellent exon, frame 1"
36392. .36663
/rpt_family="Alu"
36901. .37222
/notes="100% identity EST ou55c09.x1"
/db_xref="dbEST:A1025011"
36901. .37164
/standard_name="STS-G-9983"
/db_xref="dbSTS:G26554"
/misc_feature complement(37595. .37654)
/notes="GRAIL 2 excellent exon, frame 0"
complement(join(38063. .38218,38462. .38578,38741. .38995
39071. .39205,39532. .39630,39935. .40048,40300. .40410,
40503. .40661,41868. .41972,42103. .42225,42492. .42569,
44379. .44507))
/standard_name="histidyl-tRNA synthetase"
/notes="78%-100% protein identity GenPept:U18937"
misc_feature complement(38069. .38215)
/notes="GRAIL 2 excellent exon, frame 0"

```

Note: remainder of annotations omitted

[illegible]

```

Cp      460  TCCGGATGTGGATACACTTCGTCCCTTCATAAACAAGAACACGGAGCTTCGTCCG  401
Db      16038  CSM 16040
Cp      400  GGC 398
::

RESULT  15
LOCUS   AR024229      965 bp      DNA
DEFINITION  Sequence 22 from patent US 5795961.
ACCESSION  AR024229
NID      93977523
VERSION  AR024229.1  GI:3977523
KEYWORDS
SOURCE   Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 965)
AUTHORS  Wallace,J.,Paul, Harris,W.J., Carr,F.J., Old,L.J., Welt,S. and
          Kitamura,K.
TITLE     Recombinant human anti-Lewis b antibodies
JOURNAL   Patent: US 5795961-A 22 18-AUG-1998;
FEATURES
          Location/Qualifiers
            source          1..965
                           /organism="unknown"
BASE COUNT  192 a 170 c 226 g 205 t 172 others
ORIGIN
Query Match 4.5%; Score 32; DB 25; Length 965;
Best Local Similarity 16.9%; Pred. No. 4,96e-06;
Matches 23; Conservative 61; Mismatches 51; Indels 1; Gaps 1

Db      788  VSGGVRSTICTASDYTTISYGWVGRGWDYGGGYNYNNGKRGRVTMTDTSNSRSSVT  847
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp      656  AGAGAGAGCCCTC-AGGATCTCTGTGAGATGTCTCTCGGAAGAAGTAGATCTTGTC  598
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      848  AADRAVYVCVRGRSDSGDGYMGCTIVYASHVKMTSSSSASVGDRTYTCSSSTHG  907
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp      537  ATCGTAAGCCCTGGCTCTTGTCACGATGGTGCGCTTTATGACACTGTGGGTTCTGCATGAC  538
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      908  NGNTYVYWKGRKAKRYV  923
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp      537  AGTATCACTGCTGTAC  522

Search completed: Wed Aug 4 01:05:40 1999
Job time : 1493 secs.

```

```
Search completed: Wed Aug 4 01:05:40 1999
Job time : 1493 secs.
```

This Page Blank (uspto)


```

RESULT 2
ID N81164 standard; DNA: 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KM E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key
FT misc_feature
FT 19..69
FT /tag= a
FT /function-multiple cloning site
FT primer_bind
FT 187..204
FT /tag= b

EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
WPI: 88-279927/40.

DR 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; P: English.

CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.

SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 6.1%; Score 43; DB 1; Length 204;
Best Local Similarity 7.2%; Pred. No. 7.14e-13;
Matches 7; Conservative 56; Mismatches 34; Indels 0; Gaps 0;

DB 92 hhyrmdbnyrdyndsdaaawccyrsvkydcynachddhyvbbvynvnhnc 151
QY 320 CCCGGCACCCACGCTGCTGGACCTGGTGAATGCACGTGTGCCACCTTGGCGAGATG 379
DB 152 ncccbnhvchvbnhmrhwayrhdarddvhcv 188
QY 380 GAGGCTACGCCCTTCAGCCGCGAGACAGACGTCCC 416

RESULT 3
ID N81164 standard; DNA: 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KM E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key
FT misc_feature
FT 19..69
FT /tag= a
FT /function-multiple cloning site
FT primer_bind
FT 187..204
FT /tag= b

EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
WPI: 88-279927/40.

```

```

PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; P: English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.

SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 6.1%; Score 43; DB 1; Length 204;
Best Local Similarity 7.9%; Pred. No. 7.14e-13;
Matches 7; Conservative 52; Mismatches 30; Indels 0; Gaps 0;

DB 90 tthyrmbdnyrdyndsdaaawccyrsvkydcynachddhyvbbvynvnhn 149
CP 350 TTCACACAGTTCACAGACAGCTGGCGCTGCTGTGCACAGCCACAGCCCTCA 291
DB 150 ncccbnhvchvbnhmrhwayrhd 178
CP 290 CTCGCGCTCTCCAGAGAGATGATGATT 262

RESULT 4
ID Q51746 standard; cDNA: 91 BP.
AC Q51746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KM Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW ss.
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PF 24-MAY-1993; 108325.
PR 26-MAY-1992; US-889651.
PA (BECT) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WPI: 93-378844/48.
PT New oligo:nucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3; Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also Q51735-45 and Q51747-59.

SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 5.8%; Score 41; DB 9; Length 91;
Best Local Similarity 0.0%; Pred. No. 1.37e-11;
Matches 0; Conservative 44; Mismatches 3; Indels 0; Gaps 0;

DB 14 hsyvrvhvhshhsvhvhvsvvvhvvhvvhvvhvsv 60
QY 179 TCCCGAGGCAAGAACGATCTGTGCGCAGTGAATATGCGCTCC 225

RESULT 5
ID Q70465 standard; DNA: 114 BP.
AC Q70465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KM TSAR: totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker.

```


[illegible]

[illegible][illegible]

CC or compens. comprising a TSAR binding domain can be used in vivo to

CP 345 CAGGTTCCAGCAGCTGGGGTGCCGGGCGTTGGTGCCACAGGCCAGGCCCC 294

Search completed: Wed Aug 4 01:33:14 1999
Job time : 188 secs.

This Page Blank (uspto)

 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 (TM)

 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 3 16:28:58 1999; Maspar time 19.44 Seconds

Tabular output not generated. 968,643 million cell updates/sec

Title: >US-09-240-410-2
 Description: (1-666) from US09240410.pep
 Perfect Score: 4968
 Sequence: 1 MTPPPGCAAPSAPRARVPG.....LAASIMGLVPTLTGLVH 666

Scoring table: PAM 150
 Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot37
 1:swissprot

Statistics: Mean 51.371; Variance 79.793; scale 0.644

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	470	9.5	403	1	VA39_VACCC PROTEIN A39.	3.68e-83
2	463	9.3	441	1	VA39_VACCV PROTEIN A39.	2.07e-81
3	163	3.3	1871	1	SEX_HUMAN TRANSMEMBRANE PROTEIN	5.75e-12
4	107	2.2	533	1	ARSB_HUMAN ARYL SULFATASE B PRECUR	4.59e-02
5	110	2.2	535	1	ARSB_FELCA ARYL SULFATASE B PRECUR	1.58e-02
6	105	2.1	227	1	US08_HOWMA HYPOTHETICAL PROTEIN H	9.21e-02
7	102	2.1	275	1	NRT2_RAT T-CELL ECTO-ADP-RIBOSY	2.57e-01
8	98	2.0	141	1	TPRO_OWEFU HYPOTHETICAL PROLINE-R	9.70e-01
9	99	2.0	197	1	TNFB_RABIT LYMPHOXYN-ALPHA PREC	6.99e-01
10	99	2.0	275	1	NRT1_RAT T-CELL ECTO-ADP-RIBOSY	6.99e-01
11	99	2.0	386	1	IMPI_EBV LATENT MEMBRANE PROTEI	6.99e-01
12	98	2.0	404	1	IMPI_EBVC LATENT MEMBRANE PROTEI	9.70e-01
13	99	2.0	442	1	SPC3_RAT SYNDICAN-3 PRECURSOR (6.99e-01
14	99	2.0	460	1	EF1A_RIRE ELONGATION FACTOR 1-AL	6.99e-01
15	99	2.0	591	1	VRP2_SALTU 65 KD VIRULENCE PROTEI	6.99e-01
16	99	2.0	591	1	VRP2_SALEN 65 KD VIRULENCE PROTEI	6.99e-01
17	99	2.0	591	1	VRP2_SALEN 65 KD VIRULENCE PROTEI	6.99e-01
18	99	2.0	593	1	VRP2_SALUD 65 KD VIRULENCE PROTEI	6.99e-01
19	99	2.0	916	1	PMAL1_HJBCA PLASMA MEMBRANE ATPASE	6.99e-01
20	98	2.0	1132	1	DNB1_HSV6U MAJOR DNA-BINDING PROTE	6.99e-01
21	100	2.0	1338	1	VRG1_HUMAN VASCULAR ENDOTHELIAL G	5.02e-01
22	95	1.9	102	1	COLL_HSVSC COLLAGEN-LIKE PROTEIN	2.54e+00
23	92	1.9	214	1	RSMB_RAT SMALL NUCLEAR RIBONUCL	6.47e+00

RESULT	1	STANDARD	PRT	403 AA.	ALIGNMENTS
ID	VA39_VACCC				
AC	P21062				
DT	01-FEB-1991 (REL. 17, CREATED)				
DT	01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)				
DT	01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)				
DE	PROTEIN A39.				
GN	A39R.				
OS	VACCINIA VIRUS (STRAIN COPENHAGEN).				
OC	VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYIRIDAE; CHORDOPOXVIRINAE;				
OC	ORTHOPOXVIRUS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE: 91021027.				
RA	GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,				
RA	PAOLETTI E.,				
RT	"The complete DNA sequence of vaccinia virus.";				
RL	VIROLOGY 179:247-266(1990).				
RN	[2]				
RP	COMPLETE GENOME.				
RA	GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,				
RA	PAOLETTI E.,				
RL	VIROLOGY 179:517-563(1990).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
DR	EMBL: M35027; G335517; -				
DR	PIR: E42521; E42521.				
SQ	SEQUENCE 403 AA; 45741 MW; FE0DB9AB CRC32;				

Db	77	LVGCTNGNPKWK-IDGSDPKHGRGYAPYQNSKYLISYN-ECVLSINISK-EG-I 132	9.5%; Score 470; DB 1; Length 403;
Qy	141	LACGTNARHSCNKLNVGIVPLGEMGYAPFSDENSLVLFEGDEVYSTRKQENYGI 200	Best Local Similarity 33.5%; Pred. No. 3.68e-83;
Db	133	KRMREFDPCGYDLYTADVNIIPDG-LRGAFLVSKDGYD-KVYILFTDTIGSR--I-VK 187	Matches 86; Conservative 59; Mismatches 95; Indels 17; Gaps 16;
Qy	201	PRFRIRGE-S-ELYSTDYVQNPQFIKATIVHQDAVYDKIYFFREDNPKNPEAPLN 258	

```

Db      188  IPVAMCMLNDEGGSSLSHRMSTLTKVLEEC-DID-GRSY-R-Q-IIHSRTIKTD-ND 241
QY      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
259  VSRVAOLCRGDOGGESSLSYSKWNTEFLKAMLVCSDAATNKNNRRLDVLPPDSQOMRW 318
QY      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
242  TLLVFEFDPSPYSACITYSMTNTIKOSFSYSLKEGYTKOLPSPAPICLPAGKVSHTTF 301
QY      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
319  TRYGVESPMWNSAVCYVSLGDIDVFNRRISLKGHSSLPLNRPRCKCLPDQQPIETFT 378
Db      302  EVIEKYNVDLDIIPKLS 318
QY      :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
379  QVADRHEFAORVEPMG 395

RESULT 2
ID      VA3_VACCV STANDARD: PRT: 441 AA.
AC      P24/64:
DT      01-MAR-1992 (REL. 21, CREATED)
DI      01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DI      01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE      PROTEIN A39.
GN      A39R OR (SALLR AND SALLR).
OS      VACCINIA VIRUS (STRAIN WR).
OC      VIRUSES: DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
CC      ORTHOPOXVIRUS.
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE: 91310644.
RA      AMEGADZIE B.Y., AHN B.-Y., MOSS B.;
RT      "Identification, sequence, and expression of the gene encoding a Mr
RT      35,000 subunit of the vaccinia virus DNA-dependent RNA polymerase.",
RL      J. BIOL. CHEM. 266:13712-13718(1991).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE: 91259063.
RA      SMITH G.L., CHAN Y.S., HOWARD S.T.;
RT      "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
RT      the right inverted terminal repeat.";
RL      J. GEN. VIROL. 72:1349-1376(1991).
CC      -! CAUTION: IN REF.2 THIS SEQUENCE IS IN TWO PART (SALLR AND
CC      SALLR!) DUE TO THE PRESENCE OF FRAMESHIFT COMPARED TO THAT OF
CC      REF.1.
CC      -----
CC      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      CC the European Bioinformatics Institute. There are no restrictions on its
CC      CC use by non-profit institutions as long as their content is in no way
CC      CC modified and this statement is not removed. Usage by and for commercial
CC      CC entities requires a license agreement (see http://www.isb.ch/announce/
CC      CC or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M61187; G335798; '-'
DR      EMBL; D11079; G222726; ALT_SEQ.
DR      EMBL; X57318; G62254; '-'
DR      PTR: S29921; S29921.
SQ      SEQUENCE 441 AA; 50185 MW; 20823368 CRC32;

Query Match 9.3%; Score 463; DB 1; Length 441;
Best Local Similarity 33.9%; Pred. No. 2,07e-81;
Matches 87; Conservative 57; Mismatches 96; Indels 17; Gaps 16;

Db      115  LVCGTNNGNPKCWK-IDGSDPFKRHGRCYAPAYONSVTIIISHNG-VLSDINISK-EG-I 170
QY      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
141  LACGANHAHPSCMNIVGTVPVLGEMRGYAPSPDENSLVFEGGEVYSTIRKQRYNKI 200
Db      171  KRMRFDFPCGYDLTAANVIPKDC-LFGAEDVKDGCTD-KYILFTPTIGSKR--I-VK 225
QY      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
201  PRFRIRGE-S-ELTSDTVQMNPFIKATVIHQOAAVDKTIYYFFREDNDPKNEAPLN 258
Db      226  IPIYAOMCLNDEGGSSLSHRMSTLTKVLEEC-DID-GRSY-R-Q-IIHSRTIKTD-ND 279
QY      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
259  VSRVAOLCRGDOGGESSLSYSKWNTEFLKAMLVCSDAATNKNNRRLDQVFLPPDSQOMRW 318

```

Db	280	111YVFDPSYSLACTYSMNINIKSEFSRSTKLEGTOKLPSPASGICLPAGKVHHTTP	339
Qy	319	TRRYGFSNPMWNSACVYSLGDIDKVFRTSLKGYHSSLPNPRGKCLPDDQPIPTETP	378
Db	340	EVIEKYNVLDIDIKPLS	356
Qy	379	QVADRHPVQAQVEPMG	395
RESULT	3		
ID	SEX_HUMAN	STANDARD:	PRT; 1871 AA.
AC	P51805;		
DT	01-OCT-1996 (REL. 34, CREATED)		
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)		
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)		
DE	TRANSMEMBRANE PROTEIN SEX PRECURSOR.		
GN	SEX.		
OS	HOMO SAPIENS (HUMAN).		
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;		
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=FETAL BRAIN, SKELETAL MUSCLE, AND EMBRYO;		
RX	MEDLINE; 96149362.		
RA	MAESTRINI E., TAMAGNONE L., LONGATI P., CREMONA O., GIULIANO M.,		
RA	BIONE S., TAMANINI F., NEEL B.G., TONIOLO D., COMOLIO P.M.;		
RT	"A family of transmembrane proteins with homology to the		
RT	metheptocyte growth factor receptor.";		
RL	PROC. NATL. ACAD. SCI. U.S.A. 93:674-678(1996).		
CC	-I- FUNCTION: PUTATIVE RECEPTOR INVOLVED IN THE DEVELOPMENT OF NEURAL		
CC	AND EPITHELIAL TISSUES.		
CC	-I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).		
CC	-I- TISSUE SPECIFICITY: PREDOMINANT IN NEURAL TISSUE DURING		
CC	DEVELOPMENT (BY SIMILARITY); WIDELY EXPRESSED IN ADULT TISSUE.		
CC	-I- SIMILARITY: WITH THE EXTRACELLULAR DOMAIN OF THE MET/ROM/HGF		
CC	RECEPTORS AND THE CYTOPLASMIC DOMAIN OF SEF, NOV AND OCT.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X87852; E183847; .		
DR	MM; 300022; .		
KW	RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.		
FT	SIGNAL	1	19
FT	CHAIN	20	1871
FT	DOMAIN	20	1220
FT	TRANSMEM	1221	1241
FT	DOMAIN	1242	1871
FT	CARBOHYD	59	59
FT	CARBOHYD	548	548
FT	CARBOHYD	637	637
FT	CARBOHYD	738	738
FT	CARBOHYD	746	746
FT	CARBOHYD	1009	1009
FT	CARBOHYD	1036	1036
FT	CARBOHYD	1073	1073
FT	CARBOHYD	1115	1115
FT	CARBOHYD	1162	1162
SO	SEQUENCE	1871 AA;	207661 MW; 9A11046A CRC32;
Db	427	VVF1GTBGSGLKKRVYVGFQDAHL-YETVYPAVD--GSPILRDLLFSFDPHRIYILISEKOV	483
Qy	426	VLYITDGTGTHKHKVPEDEHSEAFNIMEIQFERRAAIOTMSIDERRKLYVSSQWEY	485

ID	RESULT	4	STANDARD:	PRT:	533 AA.
Db	484	SOLPVENTCEYOS--CAACLGSDGDCGM	510		
Oy	486	SOVPLDCEYGGGCGHGLMSRDPYCGM	513		
DT	01-APR-1990	(REL. 14, CREATED)			
DT	01-APR-1990	(REL. 14, LAST SEQUENCE UPDATE)			
DT	15-DEC-1998	(REL. 37, LAST ANNOTATION UPDATE)			
DE	ARLSULFATASE B PRECURSOR	(EC 3.1.6.12) (ASB)			
DE	4-SULFATASE	(GAS).			
CN	ARSB.				
OS	HOMO SAPIENS (HUMAN).				
OC	EDUAROTI; METAAROT; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;				
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.				
RN	[1]	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RP	MEDLINE; 90153994.				
EX	PETERS C., SCHMIDT B., ROMERSKIRCH W., RUPP K., ZUEHSDORF M.,				
RA	VINGRON M., MEYER H.E., POHLMANN R., VON FIGURA K.;				
RT	"Phylogenetic conservation of arylsulfatases. cDNA cloning and				
RT	expression of human arylsulfatase B.";				
RL	J. BIOL. CHEM. 265:3374-3381(1990).				
RN	[2]	SEQUENCE FROM N.A.			
RP	MEDLINE; 90152677.				
RA	SCHUCHMAN E.H., JACKSON C.E., DESNICK R.J.;				
RT	"Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full-				
RT	length cDNA, and regions of amino acid identity with arylsulfatases A				
RL	and C.";				
RN	GENOMICS 6:149-158(1990).				
RN	[3]	SEQUENCE FROM N.A.			
RP	MEDLINE; 93332648.				
RA	MODARESSI S., RUPP K., VON FIGURA K., PETERS C.;				
RT	"Structure of the human arylsulfatase B gene.";				
RL	BIOL. CHEM. HOPPE-SEYLER 374:327-335(1993).				
RN	[4]	SEQUENCE OF 1-104 FROM N.A.			
RP	MEDLINE; 92028992.				
EX	LITVENS T., MORRIS C.P., GIBSON G.J., BECKMANN K.R., HOPWOOD J.J.;				
RA	"Human N-acetylglucosamine-4-sulfatase: protein maturation and				
RT	isolation of genomic clones.";				
RL	BIOCHEM. INT. 24:209-215(1991).				
RN	[5]	2-AMINO-3-OXOPROPIONIC ACID MODIFICATION SITE.			
RP	MEDLINE; 95354208.				
RA	SCHMIDT B., SELMER T., INGENDOH A., VON FIGURA K.;				
RT	"A novel amino acid modification in sulfatases that is defective in				
RT	multiple sulfatase deficiency.";				
RL	CELL 82:271-278(1995).				
RN	[6]	X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).			
RP	MEDLINE; 97184692.				
EX	BOND C.S., CLEMENTS P.R., ASHBY S.J., COLLYER C.A., HARROP S.J.,				
RA	HOPWOOD J.J., GUSS J.M.;				
RT	"Structure of a human lysosomal sulfatase.";				
RT	STRUCTURE 5:277-289(1997).				
RL	[7]	VARIANTS MPS-VI ARG-117; PRO-236 AND TYR-405.			
RP	MEDLINE; 92197625.				
RA	JIN W.-D., JACKSON C.E., DESNICK R.J., SCHUCHMAN E.H.;				
RT	"Mucopolysaccharidosis type VI: identification of three mutations in				
RT	the arylsulfatase B gene of patients with the severe and mild				
RL	phenotypes provides molecular evidence for genetic heterogeneity.";				
RL	AM. J. HUM. GENET. 50:795-800(1992).				
RN	[8]	VARIANT MPS-VI GLY-137, AND VARIANT MET-376.			
RP	MEDLINE; 92042029.				
EX	WICKER G., PRILL V., BROOKS D., GIBSON G., HOPWOOD J.,				
RA	VON FIGURA K., PETERS C.;				

RA "Mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). An intermediate
 RT clinical phenotype caused by substitution of valine for glycine at
 RT position 137 of arylsulfatase B.";
 RL J. BIOL. CHEM. 266:21386-21391(1991).
 RN [9]
 RP VARIANTS MPS-VI MET-92; GLN-95; CYS-210; PRO-393 AND PRO-498.
 RX MEDLINE: 962133747.
 RA LITJENS T., BROOKS D.A., PETERS C., GIBSON G.J., HOPWOOD J.J.:
 RT "Identification, expression, and biochemical characterization of N-
 RT acetylgalactosamine-4-sulfatase mutations and relationship with
 RT clinical phenotype in MPS-VI patients".
 RL AM. J. HUM. GENET. 58:1127-1134(1996).
 RN [10]
 RP VARIANTS MPS-VI TRP-152 AND GLN-160.
 RX MEDLINE: 941711224.
 RA VOSKOBOEVA E., ISBRANDT D., VON FIGURA K., KRASNOPOLSKAYA X.,
 RA PETERS C.:
 RT "Pur novel mutant alleles of the arylsulfatase B gene in two
 RT patients with intermediate form of mucopolysaccharidosis VI
 RT (Maroteaux-Lamy syndrome)".
 RL HUM. GENET. 93:259-264(1994).
 RN [11]
 RP VARIANT MPS-IV ARG-302.
 RX VILLANI G.R.D., BALZANO N., DI NATALE P.:
 RA "Two novel mutations of the arylsulfatase B gene in two Italian
 RT patients with severe form of mucopolysaccharidosis".
 RL HUM. MUTAT. 11:410-410(1998).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 4-SULFATE GROUPS OF THE
 CC N-ACETYL-D-GLYCOSAMINE 4-SULFATE UNITS OF CHONDROITIN SULFATE
 CC AND DERMATAN SULFATE.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.
 CC -1- DISEASE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY
 CC SYNDROME, ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
 CC THIS DISEASE IS CHARACTERIZED BY THE ACCUMULATION OF DERMATAN
 CC SULFATE IN LYSOSOMES. CLINICAL FEATURES CAN INCLUDE ABNORMAL
 CC GROWTH, SHORT STATURE, STIFF JOINTS, SKELETAL MALFORMATIONS,
 CC CORNEAL CLOUDING, HEPATOSPLENOMEGALY, AND CARDIAC ABNORMALITIES.
 CC -1- A WIDE VARIATION IN CLINICAL SEVERITY IS OBSERVED.
 CC -1- DISEASE FEATURES OF METACHROMATIC LEUKODYSTROPHY AND OF
 CC MUCOPOLYSACCHARIDOSIS: IT IS CHARACTERIZED BY A DECREASED
 CC ACTIVITY OF ALL KNOWN SULFATASES. IT SEEMS TO BE CAUSED FROM THE
 CC LACK OF POST-TRANSLATIONAL MODIFICATION OF A CYSTEINE INTO 2-
 CC AMINO-3-OXOPROPIONIC ACID.
 CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on ways
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J05225; GI79077; -.
 DR EMBL: M33373; GI19030; -.
 DR EMBL: X72735; G825628; -.
 DR EMBL: X72736; G825628; -.
 DR EMBL: X72737; G825628; JOINED.
 DR EMBL: X72738; G825628; JOINED.
 DR EMBL: X72739; G825628; JOINED.
 DR EMBL: X72740; G825628; JOINED.
 DR EMBL: X72741; G825628; JOINED.
 DR EMBL: X72742; G825628; JOINED.
 DR EMBL: S57777; G236698; -.
 DR PIR: A35078; A35078.
 DR PDB: 1FSU; 04-FEB-98.
 DR MIM: 253200; -.
 DR MIM: 272300; -.
 DR PROSITE: PS00523; SULFATASE_1; 1.
 DR PROSITE: PS00149; SULFATASE_2; 1.
 DR PFM: PF00884; Sulfatase; 1.
 DR

Query Match	Similarity	Score	DB 1	Length	533
Best Local	41.3%	Pred. No. 4,59e-02			
Matches	26	Conservative	13	Mismatches	19
				Indels	5
				Gaps	
Db	4	RGASLSPRG--PGR-RLLPVPLLLLPAPGAGASRPPLVLL-ADDLGMD	59		
	8	RAAPSAPRAYVGPAPRLRLRLLLW-AAASNOGHLRSGPRIFAWKGVQDR	66		
Dd	60	VGF	62		
	67	VDF	69		
RESULT	5	STANDARD	PRT	535 AA.	
AC	P33727				
DT	01-FEB-1994 (REL. 28, CREATED)				
DT	01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)				
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)				
DE	ARLSULFATASE B PRECURSOR (EC 3.1.6.12) (ASB) (N-ACETYLGALACTOSAMINE-4-SULFATASE) (GAS).				
GN	ARSB.				
OS	FELIS SILVESTRIS CATUS (CAT).				
OC	EDUARDOTA; METAOTA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;				
OC	CARNIVORA; FISSIPEDIA; FELIDAE; FELIS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-LIVER.				
RX	MEDLINE; 93052342.				
RA	JACKSON C.E., YUHKI N., DESNICK R.J., HASKINS M.E., O'BRIEN S.J.,				
RA	SCHUCHMAN E.H.;				
RT	"Feline arylsulfatase B (ARSB): isolation and expression of the cDNA,				
RT	comparison with human ARSB, and gene localization to feline				
RT	chromosome A1."				
RL	GENOMICS 14:403-411(1992).				
CC	-1- CATALYTIC ACTIVITY: HYDROLAYSIS OF THE 4-SULFATE GROUPS OF THE				
CC	N-ACETYL-D-GALACTOSAMINE 4-SULFATE UNITS OF CHONDROITIN SULFATE				
CC	AND DERMATAN SULFATE.				
CC	-1- SUBUNIT: HOMODIMER.				
CC	-1- SUBCELLULAR LOCATION: LYSOSOMAL.				
CC	-1- DISEASE DEFECTS IN ASB ARE THE CAUSE OF MAROTEAUX-LAMY				

```

CC SYNDROME; ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
CC MPS-VI HAS BEEN DESCRIBED IN SIAMSE CATS.
CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, S48472; G258856; -.
DR PIR, A44475; A44475.
DR PROSITE, PS00523; SULFATASE_1; 1.
DR PROSITE, PS00149; SULFATASE_2; 1.
DR PFAM, PF00884; Sulfatase; 1.
DR HSSP, P15848; 1ESU.
KW HYDROLASE; SIGNAL; GLYCOPROTEIN; LYSOSOME; MUCOPOLYSACCHARIDOSIS.
FT SIGNAL 1 41
FT CHAIN 42 535
FT ACT_SITE 149 149
FT MOD_RES 93 93
FT 2-AMINO-3-OXOPROPIONIC ACID
FT (BY SIMILARITY).
FT DISULFID 119 523
FT DISULFID 123 157
FT DISULFID 183 194
FT DISULFID 407 449
FT CARBOHYD 190 190
FT CARBOHYD 281 281
FT CARBOHYD 293 293
FT CARBOHYD 428 428
FT CARBOHYD 460 460
FT SEQUENCE 535 AA; 59753 MW; 4c473698 CRC32;
SQ
Query Match 2.2%; Score 110; DB 1; Length 535;
Best Local Similarity 34.6%; Pred. No. 1,586-02;
Matches 28; Conservative 19; Mismatches 31; Indels 3; Gaps 3;
Db 4 RGAASLPKPGSPRRRLTLCVPLLRLLRLLLPSRPGAGAD-RPHLYF-VLADDLGND 61
Qy 8 RAAPAPAPARVPGPPARLG-LPLRLRLLLMAAASAOGLRSGPRIFAWKGHVGD 66
Db 62 VSFHGSNRTPHLDELAAGV 82
Qy 67 VDFQTEPHYLVFHEPGSSV 87
RESULT 6
ID US08_HCMVA STANDARD; PRT; 227 AA.
AC P09730;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN HXLf4 PRECURSOR.
GN US8.
OS HUMAN CYTOMEGALOVIRUS (STRAIN AD169).
OC VIRUSES; DSDNA VIRUSES; NO RNA STAGE; HERPESVIRIDAE;
OC BETAHERPESVIRINAE; CYTOMEGALOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87169717.
RA WESTON K., BARRELL B.G.;
RT "Sequence of the short unique region, short repeats, and part of the
RT long repeats of human cytomegalovirus.";
RL J. MOL. BIOL. 192:177-208(1986).
RN [2]
RP COMPLETE GENOME.
RX MEDLINE: 90269039.
RA CHEE M.S., BANKIER A.T., BECK S., BOHNT R., BROWN C.M., CERNY R.,
RA HORSNELL T., HUTCHISON C.A. III, KOZARIDES T., MARTIGNETTI J.A.,
RA PRADDE E., SATCHELWILL S.C., TOMLINSON P., WESTON K.M., BARRELL B.G.;
RA "Analysis of the protein-coding content of the sequence of human

```



```

RT      adenoocarcinoma cells transformed with rat T cell alloantigen Rf6.2." ;
R1      J. BIOL.CHEM. 269:9420-9423(1994) .
CC      -I- FUNCTION: HAS BOTH NAD+ GLYCOHYDROLASE AND ADP-RIBOSYLTRANSFERASE
CC      ACTIVITY (TO A LESSER EXTENT).
CC      -I- CATALYTIC ACTIVITY: NAD(+) + L-ARGININE = NICOTINAMIDE +
CC      N2-(ADP-D-RIBOSYL)-L-ARGININE.
CC      -I- CATALYTIC ACTIVITY: NAD(+) + H2O = NICOTINAMIDE + ADP-RIBOSE.
CC      -I- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC      -I- TISSUE SPECIFICITY: POSTTHYMIC T CELLS.
CC      -I- PFM: TWO DISULFIDE BONDS ARE PRESENT (PROBABLE) .
CC      -I- SIMILARITY: BELONGS TO THE ARG-SPECIFIC ADP-RIBOSYLTRANSFERASE
CC      FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as their content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; M85193; G206806; -
DR      EMBL; X99123; E257753; -
DR      EMBL; X99122; E257751; -
DR      PIR; A34866; A34866.
DR      PROSITE; PS01291; ART; 1.
DR      PFAM; PF01129; ART; 1.
DR      TRANSFERASE; GLYCOSYLTRANSFERASE; GLYCOPROTEIN; NAD; SIGNAL;
DR      K W T-CELL DIFFERENTIATION; GPI-ANCHOR.
FT      SIGNAL          1         20
FT      CHAIN           21        246
FT      PROPEP          247       275
FT                                     T-CELL, EC70-ADP-RIBOSYLTRANSFERASE 2.
FT      LIPID            246       246      HYDROPHOBIC, REMOVED DURING MATURATION
FT      ACT_SITE         209       209      (BY SIMILARITY) .
FT      SEQUENCE        275 AA; 31438 MW; DBEAB84E CRC32;    GPI-ANCHOR (BY SIMILARITY) .
                                         BY SIMILARITY.
Query Match              2.1%; Score 102; DB 1; Length 275;
Best Local Similarity 30.6%; Pred.No. 2.5/e-01;
Matches 19; Conservative 17; Mismatches 21; Indels 5; Gaps 5;

Db 194 LGVYIKES-RRPOEE-VLLPGYEYVKVARTOGNEFLDSPKRKSNYCNLYSSAGAR 251
   ||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 163 LGE-MRGTAAPSDDNSLVLEEGDEVYSTIKQEIYNGK-IFRFRIKESE-LITSDIYV 219
   ::::
Db 252 ES 253
Qy 220 QN 221

RESULT 8
AC P21260; P21261; STANDARD; PRT; 141 AA.
DT 01-MAY-1991 (REL. 18, CREATED)
DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROLINE-RICH PROTEIN (FRAGMENT) .
OS OWENIA FUSIFORMIS.
OC EUKARYOTA; METAOA; ANNELIDA; POLYCHAETA; OMENIIDAE; OMENIIDAE; OMENIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90147742.
RA BAKALARA N., COLLET J., PLANELLIS R., THOUVENY Y., FONTES M.;
RT "Presence in invertebrate genomes of sequences characterized by the
RL repetition of the triplet CpGpurine";
RD BIOCHEM. BIOPHYS. RES. COMMUN. 166:66-73(1990) .
DR PIR; A34043; A34043.
DR KW HYPOTHETICAL PROTEIN; DNA-BINDING.
FT NON_TER          1
FT DOMAIN            9         58      POLY-PRO.
FT DNA_BIND          98        116     H-T-H MOTIF (POTENTIAL) .
FT NON_TER          141

```

```

SQ      SEQUENCE          141 AA;  15745 MM;   6327FEAL CRC32;
Query Match                               2.0%; Score 98; DB 1; Length 141;
Best Local Similarity 33.3%; Pred.No. 9,70e-01;
Matches       17; Conservative         12; Mismatches 22; Indels    0; Gaps    0;

Db      39 PPPPPPPPPPPPPPPPPRRRAIHNIPLFLRFKKSYSNNHMKCGOR 89
||||| ::::: | || | : | : | : | : | : | : | : | : | : | : |
QY      3 PPPGGAAPSAPRARVPGPARLGLPRLRLLTLLMMAASAOGHLSGR 53

RESULT     9
ID      TNFB_RABBIT        STANDARD; PRT; 197 AA.
AC      PI0154;
DT      01-MAR-1989 (REL. 10, CREATED)
DT      01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT      01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE      LYMPHOTOXIN-ALPHA PRECURSOR (LT-ALPHA) (TNF-BETA).
GN      LTA OR TNFB
OS      ORYCTOLAGUS CUNICULUS (RABBIT).
OC      EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC      LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
RN      (1)
RP      SEQUENCE FROM N.A.
RX      MEDLINE: 91065534.
RA      SHAIHOW A.N., KUPRASH D.V., AZIZOV M.M., JONGENEEL C.V.,
RA      NEDOSPASOV S.A.;
RT      "Structural analysis of the rabbit TNF locus, containing the genes
RT      encoding TNF-beta (lymphotoxin) and TNF-alpha (tumor necrosis
RT      factor).";
RL      GENE 95:215-221(1990).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE: 90220566.
RA      SHAHOF A.N., KUPRASH D.V., TURETSKAYA R.L., AZIZOV M.M.,
RA      ANDEREYEVA A.V., NEDOSPASOV S.A.;
RT      "Cloning and structural analysis of genes coding for tumor necrosis
RT      factor and lymphotoxin in rabbits.";
RL      MOL. BIOL. (MOSC.) 23:1743-1750(1989).
CC      -I- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH
CC      IS CYTOTXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.
CC      -I- SUBUNIT: HETEROOTIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
CC      OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
CC      -I- SUBCELLULAR LOCATION: SECRETED.
CC      -I- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outpost at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; X55745; G297168; ALT_SEQ.
DR      EMBL; M60340; G165753; -.
DR      EMBL; M60341; G165758; -.
DR      PIR; JH0309; JH0309.
DR      PIR; PND098; PND098.
DR      PROSITE; PS00251; TNF_1; 1.
DR      PROSITE; PS50049; TNF_2; 1.
FM      PFAM; PF00229; TNF; 1.
KW      CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL.
FT      SIGNAL                     26
FT      CHAIN                      197
FT      CARBOHYD                    88
SO      SEQUENCE                197 AA; 21126 MW; 38A27DBB CRC32;

Query Match                               2.0%; Score 99; DB 1; Length 197;
Best Local Similarity 33.3%; Pred.No. 6,99e-01;
Matches       31; Conservative         28; Mismatches 22; Indels    12; Gaps    11;

```

Qy	21	PARLGLPRLRLILLIMAAASGHLGRIFAWKGVGDQD--FQD-T-EPT 76
Db	58	HLVDPSPADSLMRANTDRAFLRHGESSLNN 90
Qy	77	VLFEPDS-SSV-WVGGRKVL-FDPEGRKA 106
RESULT	10	
ID	NRT1_RAT	STANDARD: PRI: 275 AA.
AC	P17982:	
DT	01-NOV-1990 (REL. 16, CREATED)	
DT	01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)	
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)	
DE	T-CELL ECTO-ADP-RIBOSYLTRANSFERASE 1 PRECURSOR (EC 2.4.2.31) (T-CELL NAD(P)(+)-L-ARGININE ADP-RIBOSYLTRANSFERASE 1) (T-CELL MONO(ADP-RIBOYL)TRANSFERASE 1) (ALLOANTIGEN RT6.1) (T-CELL SURFACE PROTEIN RT6.1)	
GN	ART2A OR RT6-A.	
OS	RATTUS NORVEGICUS (RAT).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-LEWIS A.	
RX	MEDLINE: 90192088.	
RA	HAAG F., KOCH F., THIELE H.-G.;	
RT	"Nucleotide and deduced amino acid sequence of the rat T-cell	
RL	allantigen RT6.1";	
RL	NUCLEIC ACIDS RES. 18:1047-1047(1990).	
RN	[2]	
RP	MUTAGENESIS OF GLN-207.	
RX	MEDLINE: 96275529.	
RA	MAEHAMA T., HOSHINO S.-I., KATADA T.;	
RT	"Increase in ADP-ribosyltransferase activity of rat T lymphocyte	
RL	allantigen RT6.1 by a single amino acid mutation.";	
RL	FEBS LETT. 388:189-191(1996).	
CC	-1- FUNCTION: HAS NAD+ GLYCOHYDROLASE ACTIVITY AND EXTREMELY LOW ADP-RIBOSYLTRANSFERASE ACTIVITY.	
CC	-1- CATALYTIC ACTIVITY: NAD(+) + L-ARGININE = NICOTINAMIDE + N2-(ADP-D-RIBOSYL)-L-ARGININE.	
CC	-1- CATALYTIC ACTIVITY: NAD(+) + H(2)O = NICOTINAMIDE + ADP-RIBOSE.	
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.	
CC	-1- TISSUE SPECIFICITY: POSTMYC T CELLS.	
CC	-1- PMW: TWO DISULFIDE BONDS ARE PRESENT (PROBABLE).	
CC	-1- SIMILARITY: BELONGS TO THE ARG-SPECIFIC ADP-RIBOSYLTRANSFERASE FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL: X52082; G57168; -	
DR	EMBL: M31138; G206804; -	
DR	PIR: S08464; S08464.	
DR	PROSITE: PS01291; ART; 1.	
DR	PFAM: PF01129; ART; 1.	
FW	TRANSFERASE: GLYCOSYLTRANSFERASE; GLYCOPROTEIN; NAD; SIGNAL;	
KT	T-CELL DIFFERENTIATION; GPI-ANCHOR.	
FT	SIGNAL	1
FT	CHAIN	20
FT	PROPEP	21 245
FT		247 275
FT	LIPID	246 246
FT	ACT_SITE	209 209
FT	CARBOHYD	58 58
FT	MUTAGEN	207 207
FT	SEQUENCE	275 AA; 31388 MW; 52381A84 CRC32;
FT	ACTIVITY	
FT		T-CELL ECTO-ADP-RIBOSYLTRANSFERASE 1.
FT		HYDROPHOBIC, REMOVED DURING MATURATION
FT		(BY SIMILARITY).
FT		GPI-ANCHOR (BY SIMILARITY).
FT		BY SIMILARITY.
FT		POTENTIAL.
FT		Q->E: INCREASED ADP-RIBOSYLTRANSFERASE

Query Match 2.0%; Score 99; DB 1; Length 275;
 Best Local Similarity 30.6%; Pred. No. 6.99e-01;
 Matches 19; Conservative 17; Mismatches 21; Indels 5; Gaps 5;

Db 194 LGVIRKES-FYDPOEE-VLIPGYEVYOKVKTQGYNEIFLDSPKRKSNNCLYSSAGTR 251
 163 LGE-MRGVAPFSPDENSIVLEFGDEVYSTIRKOEYNGK-IPRFRIRKESE-LYTSPTVM 219

Db 252 ES 253
 220 QN 221

RESULT 11
 ID LMP1_EBV STANDARD; PRT; 386 AA.
 AC P03230;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE LATENT MEMBRANE PROTEIN 1 (LMP-1) (P63) (P25).
 GN BNL1.
 OS EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
 OC GAMMAHERPESVIRINAE; LYMPHOCRYPTOVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 84270667.
 RA BAKER R., BANKIER A.T., BIGGIN M.D., DEININGER P.L., FARRELL P.J.,
 RA GIBSON T.J., HARFELL G., HUDSON G.S., SATCHWELL S.C., SEGWIN C.,
 RA TUFFENL P.S., BARRELL B.G.;
 RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
 RL NATURE 310:207-211(1984).
 RN [2]
 RP IDENTIFICATION OF PROTEIN.
 RX MEDLINE: 87112999.
 RA BAICHMAL V.R., SUDGEN B.;
 RT "Posttranslational processing of an Epstein-Barr virus-encoded
 membrane protein expressed in cells transformed by Epstein-Barr
 virus.";
 RL J. VIROL. 61:866-875(1987).
 RN [3]
 RP TRANSFORMING PROPERTIES.
 RX MEDLINE: 88233656.
 RA BAICHMAL V.R., SUDGEN B.;
 RT "Transformation of Balb 3T3 cells by the BNL1-1 gene of Epstein-Barr
 virus.";
 RL ONCOGENE 2:461-467(1988).
 RN [4]
 RP PROCESSING.
 RX MEDLINE: 90112654.
 RA MOORTHY R., THORLEY-LANSON D.A.;
 RT "Processing of the Epstein-Barr virus-encoded latent membrane protein
 p63/LMP.";
 RL J. VIROL. 64:829-837(1990).
 CC - FUNCTION: THE LATENT MEMBRANE PROTEIN HAS TRANSFORMING ACTIVITY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC -----
 CC EMBL: V01555; -; NOT_ANNOTATED_CDS.
 DR DIR: A03794; OQB50.
 KM TRANSMEMBRANE: PHOSPHORYLATION; TRANSFORMING PROTEIN.
 FT DOMAIN 1 24 CYTOPLASMIC.
 FT TRANSMEM 25 44 POTENTIAL.
 FT TRANSMEM 52 72 POTENTIAL.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.

FT TRANSMEM 139 159 POTENTIAL.
 FT TRANSMEM 166 186 POTENTIAL.
 FT DOMAIN 187 386 CYTOPLASMIC.
 FT PEPTIDE 242 386 P25.
 SQ SEQUENCE 386 AA; 41982 MW; 1F91D6E3 CMC32;

Query Match 2.0%; Score 99; DB 1; Length 386;
 Best Local Similarity 45.5%; Pred. No. 6.99e-01;
 Matches 15; Conservative 9; Mismatches 7; Indels 2; Gaps 2;

Db 7 RGPGRPRPRGPPPLSSGLALLLALLLFW 39
 8 RAAPSAPRA-RVPGPARLGLPLRLRL-LILW 38

RESULT 12
 ID LMP1_EBVC STANDARD; PRT; 404 AA.
 AC P29362;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
 DE LATENT MEMBRANE PROTEIN 1 (LMP-1) (P63).
 GN BNL1.
 OS EPSTEIN-BARR VIRUS (STRAIN CAO) (HUMAN HERPESVIRUS 4).
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
 OC GAMMAHERPESVIRINAE; LYMPHOCRYPTOVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92013956.
 RA HU L.F., ZABAROVSKY E.R., CHEN F., CAO S.L., ERNBERG I., KLEIN G.,
 RA WINBERG G.;
 RT "Isolation and sequencing of the Epstein-Barr virus BNL1-1 gene
 (LMP1) from a Chinese nasopharyngeal carcinoma.";
 RL J. GEN. VIROL. 72:2399-2409(1991).
 CC - FUNCTION: THE LATENT MEMBRANE PROTEIN HAS TRANSFORMING ACTIVITY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC -----
 CC EMBL: X58140; G22938; -.
 DR DIR: J01434; LABECA.
 DR PIR: S21660; S21660.
 KM TRANSMEMBRANE: PHOSPHORYLATION; TRANSFORMING PROTEIN.
 FT DOMAIN 1 24 CYTOPLASMIC.
 FT TRANSMEM 25 44 POTENTIAL.
 FT TRANSMEM 52 72 POTENTIAL.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 139 159 POTENTIAL.
 FT TRANSMEM 166 186 POTENTIAL.
 FT DOMAIN 187 404 CYTOPLASMIC.
 SQ SEQUENCE 404 AA; 43769 MW; 154E84C3 CMC32;

Query Match 2.0%; Score 98; DB 1; Length 404;
 Best Local Similarity 50.0%; Pred. No. 9.70e-01;
 Matches 16; Conservative 8; Mismatches 6; Indels 2; Gaps 2;

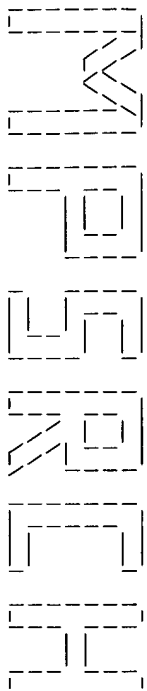
Db 8 AAPSPAPPIGPPPLSSIGLALLLALLLFW 39
 9 AAPSPAPRAVGP-PARLGLPLRLRL-LILW 38

RESULT 13
 ID SDC3_RAT STANDARD; PRT; 442 AA.
 AC P33671; P97614;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE SYNDECAN-3 PRECURSOR (N-SYNDECAN) (NEUROGLYCAN).
 GN SDC3 OR SYND3.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCITURGNATHI; MORIDAE; MORINAE; RATTUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY;
 RA KUNG C.E., DEUEL T.F.;
 RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97160597.
 RA CAREY D.J., CONNER K., ASUNDI V.K., O'MAHONY D.J., STAHL R.C.,
 RA SHOMALTER L., CIZMECI-SMITH G., HARTMAN J., ROTHBLUM L.I.;
 RT "CDNA cloning, genomic organization, and in vivo expression of rat N-
 RT syndecan.";
 RL J. BIOL. CHEM. 272:2873-2879(1997).
 RN [3]
 RP SEQUENCE OF 90-442 FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-SCHWANN CELL.
 RX MEDLINE: 92210641.
 RA CAREY D.J., EVANS D.M., STAHL R.C., ASUNDI V.K., CONNER K.J.,
 RA GARRES P., CIZMECI-SMITH G.;
 RT "Molecular cloning and characterization of N-syndecan, a novel
 RT transmembrane heparan sulfate proteoglycan.";
 RL J. CELL. BIOL. 117:191-201(1992).
 CC -1- FUNCTION: CELL SURFACE PROTEOGLYCAN THAT MAY BEAR HEPARAN SULFATE
 CC SULFATE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN NEONATAL BRAIN, HEART, AND
 CC SCHWANN CELLS, BARELY DETECTABLE IN NEONATAL OR ADULT LIVER, OR
 CC ADULT BRAIN.
 CC -1- DEVELOPMENTAL STAGE: HIGHER LEVELS IN DEVELOPING TISSUES.
 CC -1- PTM: O-GLYCOSYLATED WITHIN THE THR/SER-RICH REGION WHICH COULD
 CC INTERACT WITH LECTIN DOMAINS ON OTHER MOLECULES (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE SYNDECAN FAMILY OF INTEGRAL MEMBRANE
 CC PROTEOGLYCAN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U52825; G1399585; -;
 DR EMBL: U73184; G1657818; -;
 DR EMBL: X63143; E42404; -;
 DR PIR: S18866; S18866;
 DR PIR: A41558; A41558;
 DR PROSITE: PS00964; SYNDECAN; 1.
 DR PFAM: PF01034; Syndecan; 1.
 KW PROTEOGLYCAN; HEPARAN SULFATE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 45
 FT CHAIN 46 442
 FT DOMAIN 46 384
 FT TRANSMEM 385 409
 FT DOMAIN 410 442
 FT DOMAIN 114 302
 FT SITE 383 384
 FT CARBOHYD 78 80
 FT CARBOHYD 82 82
 FT CARBOHYD 89 89
 FT CARBOHYD 315 315
 FT CARBOHYD 367 367
 FT CARBOHYD 14 14
 FT CONFLICT 87 87
 FT CONFLICT 90 91
 FT CONFLICT 176 176
 FT CONFLICT 178 178
 A -> D (IN REF. 1).
 O -> E (IN REF. 1).
 O -> E (IN REF. 1).
 GL -> LR (IN REF. 3).
 T -> S (IN REF. 1).
 A -> D (IN REF. 1).

FT CONFLICT 184 184 A -> G (IN REF. 1).
 FT CONFLICT 238 238 T -> S (IN REF. 1).
 FT CONFLICT 270 270 E -> V (IN REF. 1).
 SQ SEQUENCE 442 AA; 45919 MW; C59F013A CRC32;
 Query Match 2.0%; Score 99; DB 1; Length 442;
 Best Local Similarity 46.8%; Pred. No. 6.99e-01;
 Matches 22; Conservative 6; Mismatches 16; Indels 3; Gaps 3;
 Db 1 MKPPRRGTAGQRYVDATHGPAR-GLLPLLLLLLAGRAAGQ 46
 1 MTPEPPGRAPASAPRA-RVP-GPARGLPLRLRLLLLMMAAASQ 45
 QY
 RESULT 14
 ID EPIA-TRIPE STANDARD; PRT; 460 AA.
 AC P34825;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE ELONGATION FACTOR 1-ALPHA (EF-1-ALPHA).
 GN TEFL.
 OS TRICHODERMA RESEI (HYPOCREA JECORINA).
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PYRENOMYCETES;
 OC HYPOCREALES; HYPOCREACEAE; HYPOCREA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OM9414 / RUTC-30;
 RX MEDLINE: 94124021.
 RA NAKARI T., ALATAIO E., PENTTILA M.;
 RT "Isolation of Trichoderma reesei genes highly expressed on glucose-
 RT containing media: characterization of the tefl gene encoding
 RT translation elongation factor 1 alpha.";
 RL GENE 136:313-318(1993).
 CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
 CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
 CC BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-TU/EF-1A SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z23012; G312887; -;
 DR PIR: S35772; S35772.
 DR PROSITE: PS00301; EFACOR_GTP; 1.
 DR PFAM: PF00009; GTP_EFTU; 1.
 DR HSSP: P07157; LAIP.
 KW ELONGATION FACTOR; PROTEIN BIOSYNTHESIS; GTP-BINDING.
 FT NP_BIND 15 22 GTP (BY SIMILARITY).
 FT NP_BIND 92 96 GTP (BY SIMILARITY).
 FT NP_BIND 154 157 GTP (BY SIMILARITY).
 SQ SEQUENCE 460 AA; 49830 MW; B9ABAB2 CRC32;
 Query Match 2.0%; Score 99; DB 1; Length 460;
 Best Local Similarity 33.3%; Pred. No. 6.99e-01;
 Matches 15; Conservative 13; Mismatches 15; Indels 2; Gaps 2;
 Db 210 YKGWEKETKAGFTG-KTLLERIDSIEPPKR-PTDKPLRLPLOVD 252
 1 YKGWEKETKAGFTG-KTLLERIDSIEPPKR-PTDKPLRLPLOVD 252
 QY 510 YCGMGQGRGICISYSSRSVLOSINAPRHKCPKPKAPLQV 554
 1 YCGMGQGRGICISYSSRSVLOSINAPRHKCPKPKAPLQV 554
 RESULT 15
 ID VRP2-SALTY STANDARD; PRT; 591 AA.
 AC P21454;
 DT 01-MAY-1991 (REL. 18, CREATED)

This Page Blank (uspto)



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 3 16:26:51 1999; Maspar time 28.08 seconds

Tabular output not generated.

Title: >US-09-240-410-2
Description: (1-666) from US09240410.pep
Perfect Score: 4968
Sequence: 1 MTPTTGGRAAPSAPRARVPG.....LAASIMLGVLPTLTGLLVH 666

Scoring table: PAM 150
Gap 11

Searched: 122810 segs, 40068593 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir60
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 50.278; Variance 90.656; scale 0.555

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	1907	38.4	653	2	T03102 semaphorin homolog A3	0.00e+00
2	737	14.8	748	2	148744 semaphorin A - mouse	8.75e-131
3	725	14.6	751	2	148748 semaphorin E - mouse	4.38e-128
4	716	14.4	749	2	G01856 semaphorin V - human	4.61e-126
5	688	13.8	753	2	G02173 semaphorin III family	8.77e-120
6	685	13.8	772	2	A49069 collapsin - chicken	4.12e-119
7	669	13.5	772	2	148747 semaphorin D - mouse	1.56e-115
8	667	13.4	666	2	158169 semaphorin III - mus	4.37e-115
9	661	13.3	771	2	D49423 semaphorin C - mouse	3.58e-104
10	618	12.4	782	2	148746 M-sema F protein prec	5.15e-83
11	522	10.5	834	2	566498 semaphorin B - mouse	7.12e-74
12	480	9.7	760	2	148745 A39R protein - vacci	1.04e-71
13	470	9.5	403	2	E42521 hypothetical protein	3.37e-70
14	463	9.3	441	2	S29921 fosciclin IV precursor	5.73e-67
15	461	9.3	730	2	JH0798 semaphorin I precursor	2.13e-54
16	448	9.0	1074	2	JC5928 semaphorin II precurs	9.90e-50
17	389	7.8	711	2	A49423 semaphorin I - fruit	1.45e-44
18	367	7.4	656	2	B49423 semaphorin II precurs	1.45e-44
19	362	7.3	724	2	C49423 Sal19R protein - vacc	7.49e-16
20	293	5.9	295	2	JC4975 plexin 3 precursor -	3.14e-11
21	196	3.9	1884	2	JC4976 plexin 1 precursor -	6.96e-11
22	170	3.4	1872	2	JC4976 plexin 1 precursor -	6.96e-11
23	168	3.4	1894	2	JC4980 plexin 1 precursor -	6.96e-11

24	171	3.4	1905	2	I51553 Plexin - African claw	2.11e-11
25	157	3.2	122	2	J01845 14R protein - variola	5.17e-09
26	157	3.2	122	2	H36852 A43R protein - variol	5.17e-09
27	155	3.1	142	2	T01776 SalP1R protein - vacc	1.12e-08
28	120	2.4	775	2	E70320 polyribonucleotide nu	3.53e-03
29	115	2.3	422	2	S32357 glial growth factor -	1.85e-02
30	107	2.2	533	1	KJH0AB N-acetylglactosamine	2.37e-01
31	110	2.2	535	2	A44475 N-acetylglactosamine	9.24e-02
32	105	2.1	227	2	S09922 hypothetical protein	4.39e-01
33	105	2.1	235	1	Q0BEC9 HXLRP protein - human	4.39e-01
34	102	2.1	275	2	A34866 T-cell surface protei	1.09e+00
35	103	2.1	406	2	B69064 conserved hypofactica	8.05e-01
36	99	2.0	197	2	JH0309 tumor necrosis factor	2.64e+00
37	99	2.0	275	2	S08464 T-cell alloanitigen RT	2.64e+00
38	99	2.0	366	1	Q0B550 latent membrane prote	2.64e+00
39	99	2.0	460	2	S35772 translaton elongatio	2.64e+00
40	99	2.0	563	2	S78224 virulence-associated	2.64e+00
41	99	2.0	591	2	S26565 virulence-associated	2.64e+00
42	99	2.0	597	2	A46050 thyroid/steroid recep	2.64e+00
43	100	2.0	1091	2	A58532 glial cell membrane g	1.97e+00
44	99	2.0	1338	2	S09982 protein-tyrosine kina	2.64e+00
45	99	2.0	1375	2	JC5148 hepatocyte growth fac	2.64e+00

ALIGNMENTS

RESULT	ENTRY	1	ALIGNMENTS
ENTRY	T03102	#type complete	
TITLE	semaphorin homolog A3 - alcelaphine herpesvirus 1		
ORGANISM	#formal_name alcelaphine herpesvirus 1		
DATE	24-Mar-1999	#sequence_revision 24-Mar-1999	#text_change
ACCESSIONS	T03102		
REFERENCE	Z1840		
#authors	Ensser, A.; Pfanz, R.; Fleckenstein, B.		
#journal	J. Virol. (1997) 71:6517-6525		
#title	Primary structure of the alcelaphine herpesvirus 1 genome.		
#accession	T03102		
##status	Preliminary; translated from GB/EMBL/DBJ		
##molecule_type	DNA		
##residues	1-653	##label ENS	
##cross-references	EMBL:AF005370; NID:g2337967; PID:g2337970		
SUMMARY	#length 653	#molecular-weight 72645	#checksum 5501
Query Match	38.4%	Score 1907;	DB 2; Length 653;
Best Local Similarity	46.2%	Pred. No. 0.00e+00;	
Matches	279;	Conservative 108;	Mismatches 203; Indels 14; Gaps 14;
Db	45	PAMGTCVCSIRLLMIT-SAITAKSRFIDKPRILVLTGFCOHRF-FGPOPHVYLFH 102	
Qy	22	PARG-LPLRLRLILLMAAASNOGHLKSGPRIFAWKCHVGODRDFQOTEPHYLFH 80	
Db	103	SLNSDDVYVGNNITVLFDFRASHSNASTALINTSTHNRHSSTCENFTLLHNQDTGL 162	
Qy	81	EPGSSVWVGKRGVYLFDFPECKNNSVFTVNGSTKSGCLDRKDCENYITLLERSEGL 140	
Db	163	LACGINSQAKRSCW-LINNLITOTFLGPKLGLAPSPSSGNNVYLDQNDYSTINLXSLG 221	
Qy	141	LACGTNARHPSCNWLVNGVYLP-LGEMRGYAPSPSPENSVLVEGDEVYSTIRK-OEYNG 198	
Db	222	S-HKFRRIAGOVETLSDTAMHRPOFVQATVAVKHNSYDOKIYFFFEQNSHSPKOPHT 280	
Qy	199	KIRPFRIKRESEHETIDTYMOPFKATIVQDQAVYDKIYFFFEQNPDKNPEPLN 258	
Db	281	VPRVGVCCSDOGGESSLSVYKWTFLKARLACVDYDTGRINVELODIFLWQAPENSWE 340	
Qy	259	VSRVACLRDGDGESSLSVSKWNTFLKALVCSDAATNKNFRLDVFLLPDSGQWRD 318	
Db	341	TLIYGLFSPWNSAVCVFVYKIDHYFKTSKANKYHHKLPTRPPOCKMNHQVPETF 400	
Qy	319	TRYGVGFSPWNTSACVYVSLGIDIKVFRTSSLKNGHSSLPNRPKGCCLPDQDPIPETF 378	
Db	401	QVADRYEVADPYQKNNNAFPIQSKYIYTKLVYRVEYG-GVFMAITFYLTIKGTIH 459	

Oy		379	OYADHNPVAAORVEPMGKLPLEHSHTYHQKAVHERMQSHGETH-VLYLTTRGTTH	437
Db		460	IYVRNEDSNSTTANINEPFOKPAPIONILLDNTNLKLXVSEMEVEVLDCSYVG	519
Oy		438	KVERGEQEHSAFNIMIEIOEFRAAALIQMSIDAERRKLVSQSQVEQVPLDCEYGG	497
Db		520	NDFECFMSRPLCTWYNIC-S-FK-QRYSVETGCFANFTISEMCGDHAFPTVVKHOVS	576
Oy		498	GGCCHCLLSRPDYGMWDGRCSISYSSEERVSLOINPAEPH-KE-CPNPKPRAPLORKS	555
Db		577	IPILNSVSLSPAYSNHADYFWMTGDTEKRCHVKTHKNDICLLIANSTAINGFHYOCM	636
Oy		556	LAPNSRYTLSCPMESRAHTYSWRKHEVNESCPEGHOSPCLIFIENTLAQQYIGHFCDA	615
Db		637	KEDS 640	
Oy		616	QEGS 619	
RESULT	2			
ENTRY	148744	#type complete		
TITLE	semaphorin A - mouse			
ORGANISM	#formal name Mus musculus #common_name house mouse			
DATE	02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 04-Sep-1998			
ACCESSIONS	148744			
REFERENCE	148744			
#authors	Puschel, A.W.; Adams, R.H.; Betz, H.			
#journal	Neuron (1995) 14:941-948			
#title	Murine semaphorin D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension.			
#cross-references	MUID:95267431			
#accession	148744			
#status	preliminary; translated from GB/EMBL/DDAJ			
#molecule_type	mRNA			
#residues	1-748 #label RES			
##cross-references	EMBL:X85990; NID:g854323; PID:g854324			
GENETICS				
#gene	sema			
CLASSIFICATION	#superfamily semaphorin			
SUMMARY	#length 748 #molecular_weight 82894 #checksum 9017			
Query Match	14.8%; Score 737; DB 2; Length 748;			
Best Local Similarity	33.7%; Pred. No. 8,75e-131;			
Matches 168; Conservative 112; Mismatches 166; Indels 52; Gaps 36;				
Db		112	ECAFNFVRLAHYNNTHLLACTGAHFHTCALMKMATAGTAHGTAHGPEKLEDGKGTPTDP	171
Oy		125	DCENTITLLERSRG-LLACGTANARHSC---WNLVNT-V-V-P-LGEMRGASPPSP	174
Db		172	RHRPPSVLYGEELISGV-TADLMGRDPTFRSIGONPSLRTPHPDSRWMLNEKFYKVM	230
Oy		175	DENSLYLFEDEGYSTRIRKOENYK-IPIRRRIKGSELTY-S-DIV-MQNPFIKAT-I	229
Db		231	PESENPDODDKIYFFRESAVEAAPAMGRMSVSFVGOICRNLDGGORSI-VNKRTTFELKAR	289
Oy		230	VHQOAYXDRIKYFFREDVDNDPKME-APLVWSVAQLDGGDGGSLSVSKMNFLKAM	288
Db		290	LYCVSPVEGECTNHFDQADVFLIS-SR-DROTPLLIAVFSTSSGVFGQSAVCVSYMDEV	347
Oy		289	LVCS--DAATTKKNNRLODVFLLPDSGGQRDRYGVGSNPMW-Y-SAVCVYSTIGDID	343
Db		348	RAFGLPDKHSGPRHNQWVSYGSRVPDRPMCCKTFEFSSTKDPPDDVIYOFGNHHFLM	407
Oy		344	KVFRTS-SLK-G-----YHSSLPNRPEKCLPDO-QPIP-TETF--QVAD-R-HPEV	387
Db		408	YNPLPLPMG-G-R-PLFLGVAGYFTTOIADRVAAAGH-YDLVLTGSTDVGYLKIYSPK	464
Oy		388	AQRVEPMGPLKTPLE-H-SKHYYOKVAVARMQASHGELFHVLYLTTRGTTHKKAYE-P-	442
Db		465	GRPNSEGILLLEEDLFVEDSAITSMQISSKROOLTVVASRAAVATAIHRCITALGRCAE	524

QY	443	GEQEHSAFNFNTMIQIPRRRAAIAIOTNSLDAERKLIYSSQWEVSOQVPLDLCEVYGGCGG	502
Db	5325	CCLARDPACAMDGACCTAFQPTAKRRFRRODINRGDPSTLCSGSSSHVLEKUYGVES	584
QY	503	CLMSDPPYCGMDQGRCSISYSSBSVLT--QSINAPRHHCECPKPRDKAPLQKSLAPNS	560
Db	585	GSATFCEPBRSLQAHVOM	602
QY	561	RYLSCPMESRHAATYSW	577
RESULT	3		
ENTRY	148748	#type complete	
TITLE	semaphorin E - mouse		
ORGANISM	#format_name Mus musculus #common_name house mouse		
DATE	02-Jul-1996	#sequence_revision 02-Jul-1996	#text_change 04-Sep-1998
REFERENCES	148748		
REFERENCE	148744	Puschel, A.W.; Adams, R.H.; Betz, H.	
#authors	Neuron (1995) 14:941-948		
#journal	Murine semaphorin D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension		
#title	Murine semaphorin D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension		
#cross-references	MUJID:95267431		
#accession	148748		
#status	preliminary; translated from GB/EMBL/DBJ		
#molecule_type	mRNA		
#residues	1-751	#label RES	
#cross-references	EMBL:X85994; NID:9854331; PID:9854332		
GENETICS			
#gene	semd		
CLASSIFICATION	#superfamily semaphorin		
SUMMARY	#length 751 #molecular_weight 85259 #checksum 8961		
Query Match	14.68; Score 725; DB 2; Length 751;		
Best Local Similarity	33.08; Pred. No. 4.38e-128;		
Matches 148; Conservative 110; Mismatches 149; Indels 41; Gaps 28;			
Db	167	SENPVNTVSVMINEBELSGMYI-DFMGDTAIFRSILTRMQLRTDOHNSKWLSEPMFVD	225
QY	171	PFSDENSLVFEDDEVYSTRKQETNGK-IPFRRIRESSELYTS--DTV-MQNPFQIK	226
Db	226	AHVIPDGTDPNADKYFEFFERLTDNNRSYKQISHMIAKICPNDTGGQKSL-VNKKTFPL	284
QY	227	ATVHQD-QAVYDXYFEFREDNPKNPEAPLVNSVQAQLCGDGGGESSLSVSKMNTFL	285
Db	285	KARLYCSTYDDEGETHFDLEDEYLL-ETDNP-RITLYIGYFTTSSVYKGSANVCYHL	342
QY	286	KAMLYCS--DA-ATNKNFNRLQVLFPLPDSGQWBRTRYGVFNSPMWN-Y--SAVCYSL	339
Db	343	SDIOGVFNGPAHKEGNNHQLISYOGRIYPYRPGTQPGGAFIPNNRTTKDPDDVYFIR	402
QY	340	GDIDKIVRTS-SLK-G-----YHSLRPKPRGCKLPQO-QP-I-PIETF--QVADNRP	385
Db	403	NHPLMYSISPIHRRLIVRIGTDYKYTKIADRVNAADG-RYHVLFGTDGRTQVQYVV	461
QY	386	EVAQRVEPMGL-KTPLF-H--SKYHOKAYAHMQASHGEFHHVLYLTTDGTGTHKAYE	441
Db	462	LPTNYSASGELILELEVEFNHVPITTTMEISSKQOOLYSSNEGVSOVSLARCHITYTAC	521
QY	442	-PGDEHSFAFNIMEIOPFRRAAIIOTMSLDAERKLIYSSQWEVSOQVPLDLCEVYGGCG	500
Db	522	ADCLLADPACAMDGHGCSRFYTGKRRSRQDVRKGNLITGCRGFNLKATRNAAEITYQ	581
QY	501	HGCLMSADPPCGWQGCISYSS--RSVLDSINPAEHRKECP--NRPDPAPLQKYSL	556
Db	582	GVRNNTSTLECAPKSPQASIKMLQDKD	609
QY	557	A-PNSRYTLGCPMESRHAATYSW-RKEN	582
RESULT	4	G01856	#type complete
ENTRY			

TITLE	semaphorin V - human
ORGANISM	#formal_name Homo sapiens #common_name man
DATE	21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 04-Sep-1998
ACCESSIONS	G01856
REFERENCE	G08634
#authors	Sekido, Y.
#submission	submitted to the EMBL Data Library, June 1995
#accession	G01856
#status	preliminary; translated from GB/EMBL/DBJ
#molecule_type	mRNA
#residues	1-749 #label SEQ
#cross-references	EMBL:U28369; NID:g974283; PID:g974284
CLASSIFICATION	#superfamily semaphorin
SUMMARY	#length 749 #molecular_weight 83121 #checksum 2747
Query Match	14.4%; Score 716; DB 2; Length 749;
Best Local Similarity	33.0%; Pred. No. 4,616-146;
Matches	167; Conservative 115; Mismatches 170; Indels 54; Gaps 411
Db	112 ECOMFVLLHAYNTHLLACGTGAHPICAFVEVGHRAEPEVLRDPERIDTGKSPYD 171
Oy	125 DCENITTLER-RSEGLACGTNARHPSC-MNLV-N-G-TVPDL-G-EM-KGAPFS 173
Db	172 PRHRAASVVGEEELYSQA-ADLMQRDTTIRSLGQRSLTEPHDSKWLNEPFAVFW 230
Oy	174 PDENSLVLEEGDEYVSTRKQENYK-TPRRRIIGSESELYT-S-DIV-MONPQFIAT- 228
Db	231 IPESFNPDQDKYFFRETAEVAEPALGRLSVSRVQICRNDVGGORS-LVNKMTFLKA 289
Oy	229 IVHQDAIDDKIYIFFRDNPDKNE-APLNVSRVAQICRDOGESLSYSKNMTFLKA 287
Db	290 RLVCSPVEGDTHFDQIDVFLLS-SRDH-RTPLVAVFSTSSSIFQGSAYCVYSMDV 347
Oy	288 MLVCS--DAATNKNFNRLQDVFLLPDPGQMRDTRVYGVFSNPNW-N-Y-SAVCYSLGDI 342
Db	348 RRAFLGRPAHNEGPRHONVSYOGRRVYRRPQMCPSKTTGTSSTSKDFPDVDVIOGARHPL 407
Oy	343 DK-V--F--RTSS-L-K--GYHSSLJPNRPCKLPDQ-OPIT-TETF--QVAD--R-HPE 386
Db	408 MYNSVLPFGG-R-PLFLGVANNTYTTQIAARVAAADGH-ADVFIFIGDVTFLKISVP 464
Oy	387 VAQRVEPMPGLKTPF-H-SKYHKQKAAVHRMAQASHETHVLYLTDRGTHIKVYE-P 442
Db	465 KGSRPASGELLLELHVEDSAAVTSKQISSKRHOLYVASSAVAQAIALHRCAGRYCT 524
Oy	443 -GEQHSAPFNIMEIQPRRAAIAQTMSLDARRLKLYSSQMEVSYQVYLDICEVYGGCH 501
Db	525 ECLCLARDYICAMGVACTIRFQPSAKRRRRDVRNGDSTLCSGDSKRPALLEKRYGVE 584
Oy	502 GCLMRDYPGCMDOCRCSISYSERSYL--OSINPAEPHKCCPPKPDKAPL-QKVSILAP 558
Db	585 GSAFLCEPRSLARVEMTFORAGV 610
Oy	559 NSRYLSCPMBESRRATYSWR-HKENV 583
RESULT	5
ENTRY	G02173 #type complete
TITLE	semaphorin III family homolog - human
ORGANISM	#formal_name Homo sapiens #common_name man
DATE	21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 04-Sep-1998
ACCESSIONS	G02173
REFERENCE	G09275
#authors	Naylor, S.
#submission	submitted to the EMBL Data Library, October 1995
#accession	G02173
#status	preliminary; translated from GB/EMBL/DBJ
#molecule_type	mRNA
#residues	1-753 #label NAY
#cross-references	EMBL:U28276; NID:g1061350; PID:g1061351
CLASSIFICATION	#superfamily semaphorin

Query Match	13.8%	Score 688:	DB 2:	Length 753:
Best Local Similarity	31.4%	Pred. No. 8,77e-120:		
Matches 199:	Conservative 121:	Mismatches 140:	Indels 42:	Gaps 30:
SUMMARY #length 753 #molecular-weight 84941 #checksum 5681				
Db	171	PFYDKLDTASALINEELVAGYVI-DFMGTDAIRFTGKOTAMRTDOYNSRLMNDPSFIH	229	
Qy	171	PFSDENSLVFEDEYDYSTIRKEYNKG-IPRRRIIRGSELYTS--DIV-MQNPFQIK	226	
Db	230	AELIPDSEANDKLYFEFFERSAS-APQSPAVYARIGRICLNDGHCIL-VNKMSTFLK	287	
Qy	227	ATVHQDQAYDCKITYEFPREDPNKNEAPLANSRAVALCRGDGSGESSLSVSKWTFLK	286	
Db	288	ARLYCSVPGEGEITHEDELQDVRV-QQTODV-ANPIYAVFTSSGSVFSGSANVCYSMA	345	
Qy	287	AMLYCS--DATAKKNRDLQDVLDPDPSQMDWIDVYGFNSMP-WNY--SACVYSLG	340	
Db	346	DIRVFNCPFAHKRGSPNYOMPFSGKKMPYPGTCPGGTFEPSSKSTKYDPDEVINFMS	405	
Qy	341	DIDKVFIRS-SLK-G--YH-----SS-LPNRPQGLCDQ-QP-I-PIETF--QVAD--R	383	
Db	406	HPLMYQAVYPLQRRPLVYRT-GAEYRLTTIAVDQVDSADG-RYEVLFLGDRGTQVRVY	463	
Qy	384	HPVAAQRRPEPG--PLKTRPLSHSYHYQKAVAHMQASHGETFVLYLTDRGTHIKHVE	441	
Db	464	LPKDDQEMEELMEVEYFKDPAPVYKMTITSSKQOLYVASGVYTHLSLRCAQYAC	523	
Qy	442	-PGEOEHFAFINIMEIDPFRRRAALIQMTSLDAERKKLYVSSOWSEVOVPLDLCVYGGC	500	
Db	524	ADCCLRAPYCAMGQACSRRTASKRRSRQDVRHGNPIROCGFGFNSNANKNVEVOY	583	
Qy	501	HGCLMSRDPYCGMQGRGIS-TYSER-SYLSQINPDEPHKECP--NPKDPKAPLQKYSL	556	
Db	584	GVAGSAAFLECCOPRSPQATVKW	605	
Qy	557	A-PNSRYTLCSPMESRHAITYSM	577	
RESULT	6			
ENTRY	A49069	#type complete		
TITLE		collapsin - chicken		
ORGANISM		#formal_name Gallus gallus		
DATE	07-Apr-1994	#sequence_revision 07-Apr-1994		
ACCESSIONS	A49069			
REFERENCE	A49069			
#authors	Luo, Y.; Raible, D.; Raper, J. A.			
#journal	Cell (1993) 75:217-227			
#title	Collapsin: a protein in brain that induces the collapse and paralysis of neuronal growth cones.			
#accession	A49069			
#status		preliminary: not compared with conceptual translation		
#molecule_type	mRNA			
#residues	1-772	#label LDO		
#cross-references	GB:U02526; NID:9410078; PID:9410079			
CLASSIFICATION	#superfamily semaphorin			
SUMMARY	#length 772 #molecular_weight 88867 #checksum 9712			
Query Match	13.8%	Score 685:	DB 2:	Length 772:
Best Local Similarity	30.3%	Pred. No. 4.12e-119:		
Matches 190:	Conservative 158:	Mismatches 203:	Indels 77:	Gaps 60:
Db	6	GIAL-LSGLVL-AGRVACQ-HVKNVPRKLSTYKEMLESNNIVNGLANSSSYHTFL	62	
Qy	26	GLPLRLRLILLMAAASAOCHLRSG-PRIFAVMKGVH-GQDRVDF-G-QTEP-HYLVF	79	
Db	63	DE-ERSRLYVGAKDHISEFNLVNIKEYOKIYWPVSH-SRRDECKWAGKDILRECANIKY	120	
Qy	80	HEPGSSVWVGGRGRKVVYLFDEPECKN-AS-VRTYINIGSTKSC--LDK---RDENYITL	132	
Db	121	LKTYNQTLVACGGAGFHPMCTYLEVGSHPEDNIFRRHDSHFEGRGRKSPYDPLLTASL	180	

Qy	133	LERRRSG-LLACGTNARRHSC-W-NL-VN--GTVVPL--G--EM-RGYAPSPDENSLVL	181
Db	181	LVDGELVSGTA-ADFMGRDFAIFRTLGHHNPIREQDHSRWLANDPFRISAHLIPESNP	239
Qy	182	FGDEVEYSTRKOEYNGK-IPFRRIIRBESLYTS--DTV-MONPOFIKATIVHQ-DQAT	236
Db	240	DKKIYFFFRFNALIDSEHNGKATHAIRIGCIKNDGGRSL-VNKNITFLKARLICVGP	298
Qy	237	DKKIYFFREDNDKNPEAPLWVSRYVALDCKGDOGGESSLVSXKMTFLKMLVCS--DA	294
Db	299	NCIDTHFDELDOVFLM-NSKDP-KNPIYGVFTTSSNIFKSAACMCSMTDVRVFLGPY	356
Qy	295	-ATKNENRLODVFLPDPSCGMRTRYGVSPNPN-Y--SACVYSLDIDIKVFRTS-	349
Db	357	AHRDPPNOWNPYOGRVPEPRPGICPSKTFEGGFSTKDLDPVEYTFPASHNPNVFP	416
Qy	350	SLK-G-YH--S-S-LNPNRPGKLPDQ-QPI-PTETF--QV-A-DR-HPEVAQRVEPM	394
Db	417	NS-REIMAKTIDVIOFTIIVDVRDAEDGQ-YDVMTGTDIGVLKVSIPKSTHELE	474
Qy	395	GELKPLPLHPSK--YHYQVAAHRRMQASHGEFFHVLTLTRGTIHKVE-PGEQHSFA	451
Db	475	VLEEMTAFREPEYVISAAKISTKQOOLYIGATVCSOLPLRHCDYVKAACECLADPY	534
Qy	452	NIME-IOGFRRAAIIQTMSLDERKRLKLYVSQMEVSQVPLDLEVEYGGGCHGLMSRDP	510
Db	535	CAMDGSSCSRYEPTAKRRTRRDIIRNGDPLRHCSDLOHNDNPSQGLLEEKIYGVENSST	594
Qy	511	CGMDGRCISYISSE-RSVL-QSINPAEPHKECPN-PRPKAP---LQ-KVSLA-PNSRY	562
Db	595	FLECSPKSORAIV-WQFOKQNDHKVE	621
Qy	563	YLSC-PMESRHATYSWR-HKENVDSCE	588
RESULT	7		
ENTRY	148747	#type complete	
TITLE	semaphorin D - mouse		
ORGANISM	#format_name Mus musculus #common_name house mouse		
DATE	02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 04-Sep-1996		
ACCESSIONS	148747		
REFERENCE	148744		
authors	Puschel, A.W.; Adams, R.H.; Betz, H.		
#journal	Neuron (1995) 14:941-948		
#title	Mutine semaphorin D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension		
#cross-references	MIMD:95267431		
#accession	I48747		
##status	preliminary; translated from GB/EMBL/DBJ		
##molecule_type	mRNA		
##residues	1-772 #label RES		
##cross-references	EMBL:X85993; NID:98543329; PID:9854330		
GENETICS			
#gene	semd		
CLASSIFICATION	#superfamily semaphorin		
SUMMARY	#length 772 #molecular_weight 88710 #checksum 1776		
Query Match	13.5%: Score 669; DB 2; Length 772;		
Best local similarity	30.2%: Prod. No.1.56e-115;		
Matches	170; Conservative 143; Mismatches 182; Indels 67; Gaps 51;		
Db	58	HTFLPDE-ERSRLYVGAKDHIFSFLNVLNIDKFOKIWPVSY-TRRDECKWAGKIDLYECA	115
Qy	75	HTVLFHEPGSSVWVGGRGVYLFDEPECKN-AS-VRITNISTGSGC-LDK---RCE	127
Db	116	NFKIVLEANYQTHLYACGTAHPITCTTYLEVGHHPEDNIFKLODSHFENGSRGSPYDKL	175
Qy	128	NYTTLERREG-LLACGTNARRHSCWNL-V-N-GTVVPL--G--EM-RGYAPSPDE	176
Db	176	L7ASLIDGLVSGTA-ADFMGRDFAIFRTLDHNPIREQDHSRWLANDPFRISAHLIPE	234
Qy	177	NSLVLEPGDEYVITIKOEYNGK-IPFRRIIRBESLYTS--DTV-MONPOFIKATIVHQ	232

Db	235	SUNPDDKXVFFFRRENAIGESHGKATTHARIGQICNDKDFGHRSL-VNKKWTFELKARLIC	293
QY	233	-DOAYDQKITYFFREDNPDKNPAPLANSVAOLCRGDGEGSSLSVSKWMTFLKAMLYC	291
Db	294	SVGPNGIDHFEFLDQVFLM-NSKOP-KNPIYGVFTSSNIFKGSVAWCMSDVARV	351
QY	292	S--DA-ATKNFNRLDQVFLLEDPSQGMQDTRYGVFSNPMN-Y--SAVCYSIGDIDKV	345
Db	352	FLGPYAHRDGPNQWVPYOGRVPRPCTSPSKTEGGEFDSTKDLDPDITFGSRHPAMYN	411
QY	346	FRTS-SLK-G-YH--S-S-LPNRPQKCLDQ-QPI-PRTEF--QY-A-DR-HPEVAQ	389
Db	412	PVPPIIN-PRIMKIDVYNQFQIYVDRDAEDGO-YDMFEGIDVGVTLKVSVPKETW	469
QY	390	RVEPMPLTPLFHS--KHYQKVAVHNRQASHGETFVYLLTDRGTIHKVEY-PGEQE	446
Db	470	HDEEVLLEEMVYFREPRTISAMELSTKQOOLYISTSGVQOLPRHRCIDYKAKACECL	529
QY	447	HSFAEYIME-IQPFRRALAIQIMSLDAERKLIYSSQMEVSOVPLDCEYVGGGCHCGLM	505
Db	530	ARDPYCAMDGSSCSRFFPKRRTRRQDINRGDPLTHCSDELDHNNHSPLEERITYG	589
QY	506	SRDPYCGMDQGRCIYSYSE-RSVL-QSILNAPRHKPEPN-PRDK-AP-LQ-KVSLA-	557
Db	590	ENNSTFLECSPKSORALYV-WQ 610	
QY	558	PNSRYILSC-PMESRRATYSWR 578	
RESULT	8		
ENTRY	158169	#type fragment	
TITLE	semaphorin III - mouse (fragment)		
ORGANISM	Journal_name Mus musculus #common_name house mouse		
DATE	26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 04-Sep-1998		
ACCESSIONS	158169		
REFERENCE	158169		
#authors	Messersmith, E.K.; Leonardo, E.D.; Shatz, C.J.;		
#journal	Testesler-Lavigne, M.; Goodman, C.S.; Kolodkin, A.L.		
#title	Neuron (1995) 14:949-959		
#cross-references	Semaphorin III can function as a selective chemorepellent to pattern sensory projections in the spinal cord.		
#accession	MU1D:95267432		
#status	158169		
#molecule_type	preliminary: translated from GB/EMBL/DBJ		
#residues	1-666 #label RES		
#cross-references	GB:140484; NID:g703189; PID:g703190		
GENETICS			
CLASSIFICATION	SemallII		
SUMMARY	#superfamily semaphorin		
	#length 666 #checksum 9654		
Query Match	13.4%: Score 667; DB 2; Length 666;		
Best Local Similarity	30.1%: Pred. No. 4.37e-115;		
Matches	150; Conservative 130; Mismatches 168; Indels 50; Gaps 39;		
Db	60	RKSPDYDKLLTASLLIDGELYSGTA-ANFMGRDFAIFRTLGHHNPIRTEOHDSRWLPND	118
QY	167	RGYAPFSPDENSLVLEFGDEVYITRKQETNGK-IPFRIRIRSESLYYS--DTV-MQNP	222
Db	119	RFISAHLLPEDNEDDKVYFFFRRENAIDGESHGKATTHARIGQICNDKDFGHRSL-VNKKW	177
QY	223	QFIKATVHQ-DOAYDQKITYFFREDNPDKNPAPLANSVAOLCRGDGEGSSLSVSKW	281
Db	178	TFPKARLICVPPNGIDHFEFLDQVFLM-NSKOP-KNPIYGVFTSSNIFKGSVAWCMSDVARV	235
QY	282	NTFLKAMLYC--DA-ATKNFNRLDQVFLLEDPSQGMQDTRYGVFSNPMN-Y--SAVC	335
Db	236	MYNSMDVARVLLGPYAHRDGPNQWVPYOGRVPRPCTSPSKTEGGEFDSTKDLDPDIT	295
QY	336	VYSIGDIDKV----F--RTSS-IR--GYHSILNRPKPKCTPNO-ORT-PRTEF--QY-A-291	

Db	296	FASHPAPYVNPVPLNN-RPIMITDVYQSTQIVNVRVDAEDGQ-VDVMEIGDVGTVL	353
Oy	382	-DR-HPEVAQVBERGPKLTPFHS--KYHQAQVAMHMOASHETFNVLVLTDRGTIH	437
Db	354	KVSVAPKETWMDLEVLLEEMTVPEPTITAMELSTKOQOLYIGTAGVACLPHACDI	413
Oy	438	KVVE-PGEOHSAFNIME-IQPFRRAAIQTMSLDAERKRLTVSSQMEVSOVPLDCEV	495
Db	414	YKACAEECLARDPYCAWDGSSCSRYEFTPAKRTRRODIRNGDPLTHCSDLOHHNHHG	473
Oy	496	YGGCGHCLMSRDPYCGWDGRCISYSS-RSVL-QSINPAEPHKECPPN-KPPDK--AP	550
Db	414	SLEETIITGVENSSFFELCSKSORALYYWOPORNRDKEIKMGHIIITEGGLLRS	533
Oy	551	-LQ-KVSLA-PSNRYTLSC-PMESRHATY-SW-RHKENVEGSCBPG-HQ-SPNCILFIEN	602
Db	534	LQKQDNGVYLCHAVEHGF	551
Oy	603	LTAQOYGHYFCEADGCGT	620
RESULT	9		
ENTRY	D49423	#type complete	
TITLE	Semaphorin III precursor - human		
ORGANISM	Homio sapiens	#common_name man	
DATE	06-Jan-1995	#sequence_revision 06-Jan-1995	#text_change
ACCESSIONS	D49423		
REFERENCE	D49423		
authors	Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.		
#journal	Cell (1993) 75:1389-1399		
#title	The Semaphorin genes encode a family of transmembrane and secreted growth cone guidance molecules.		
#accession	D49423		
#status	preliminary; nucleic acid sequence not shown		
#molecule_type	mRNA		
#residues	1-771	#label KOL	
GENETICS	#cros-references GB:126081; NID:9799328; PID:9436560		
#gene	GDB:SEMA1		
CLASSIFICATION	#cros-references GDB:283448		
SUMMARY	#superfamily semaphorin		
	#length 771	#molecular_weight 88889	#checksum 6249
Query Match	13.3%	Score 661;	DB 2; length 771;
Best Local Similarity	30.3%	Pred. No. 9.55e-114;	
Matches 151;	Conservative 126;	Mismatches 174;	Indels 47; Gaps 37;
Db	166	RKSPYDPKLLTASLLIDGELYSGRN-ADFMGRFAFRTLGHHNPIRTEDHDSRWLNDP	224
Oy	167	RGVAFSPDENSIVLFESEDEVYSTRIRGEYNGK-IPFRFRIRGSESLYTS-DIV-MQNP	222
Db	225	KFISAHLISSEDNBPBDKVYFFPEMNAIDGSHSKATHARIGQICAKNDFGCHSL-VNKK	283
Oy	223	QFIKATIVHQ-DOAYVDKDIYFFEDNDKMPKAPLVNSVRAOICRCDGEGESSLSYSKW	281
Db	284	TTFLKARLCSVPNGNGIDTHEDLOVEFM-NKKDP-KNPVVYGVYTTSSNIFKSAVYC	341
Oy	282	NTFLKAMLYCS-DA-ATNKNFNRLQDFLLPDRDSGGWRDTRVYGVSNPNV-Y-SAVC	335
Db	342	MYSMSDVARVYVLPAGYAHNRDGPYNOVMVPOGVAPYPRPGTGPSTKFGGFDSTKDLPPDVIT	401
Oy	336	VYSLGDDIDKVERTS-SLK-G-YH-S-S-LPMPRPKGKLPQ-QPI-PLTEF-QV-A	381
Db	402	FARSHAPYVNPVFPNNRPIYIKTDVNTQFTQIVVDRVDAEDGQ-YVMFETIGDVGTVL	460
Oy	382	-DR-HPEVAQVBERGPKLTPFHSKYHQAQVAMHMOASHGTFEFLVLTDRGTIHK	438
Db	461	VVSPKKEWMDLEVLLEEMTVPEPTITAMELSTKOQOLYIGTAGVACLPHACDI	520
Oy	439	VVE-PGEOHSAFNIME-IQPFRRAAIQTMSLDAERKRLTVSSQMEVSOVPLDCEV	496

Db	521	GKADECLLRDPVCAWDSACGRYPPLAKRRTRRRDINDGNPRLHCSPLDHHNDHNGSP	580
Oy	497	GGGCHGLMSDDPCGMDGRCISYSSSE-RSYL-OSINPAEPHKECPKPKDC---AP	550
Db	581	EERIIYEVENSITFLECPSPKQALVYWOFORNEERKEEIHVDHIITFDGGLLRISIQ	640
Oy	551	LQKXSLA-PNSRYTLC-S-PMESHARY-SW-RIKENV-EQSEPGHQ-SPNCLITENLT	604
Db	641	QKDSGNYLCHAVEHGFIO	658
Oy	605	AQYGHYCFCEAQEGSYFR	622
RESULT	10		
ENTRY	148746	#type fragment	
TITLE	semaphorin C - mouse (fragment)		
ORGANISM	#formal name Mus musculus #common name house mouse		
DATE	02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 28-Feb-1997		
ACCESSIONS	148746		
REFERENCE	148744		
#authors	Puschel, A.W.; Adams, R.H.; Betz, H.		
#journal	Neuron (1995) 14:941-948		
#title	Mutine semaphorin D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension		
#cross-references	MUOJ:95267431		
#accession	148746		
#status	preliminary; translated from GB/EMBL/DBJ		
#molecule_type	mRNA		
#residues	1782 #label RES		
#cross-references	EMBL:X85992; NID:9854327; PID:9854328		
GENETICS			
#gene	semc		
SUMMARY	#length 782 #checksum 1571		
Query Match	12.4%;	Score 618; DB 2; Length 782;	
Best Local Similarity	33.7%;	Pred. No. 3,586-104;	
Matches 149; Conservative	96; Mismatches 148; Indels 49; Gaps 32;		
Db	74	KRDQNTIKILPLNSHLLTGGTAFLPCATIHIASTFLNADGANYILEDGCHOPF	133
Oy	123	KRDENIT-LLERRSGGLAGCTNARHPSG-W-N---LV-NGTV-VPLGEMGYAPF	172
Db	134	DPNFSTALVYDGLYGTVS-SFGQNDPAISROSSRPTKESSLNMLDPAFASATS	192
Oy	173	SPDENSILVLEGGDEVYSTIRKQETKNKIRFRIRGESLITSDYV--MGNQFKATIT	230
Db	193	PESIGSDIGDDKIYFFSEFGGEFFENT1-VSRVAVCKGDEGERVLD-QRMISFL	250
Oy	231	HQDQ-A-Y--DKRIYEFREDNDP-KNPEAPLNVSVQAQLCGDGGESLSVKWNFL	285
Db	251	KAOLCGRPDGPPENVLDVFTLNPPQ-DNRKITLSIVFTSQMHRGTTESALCVETM	309
Oy	286	KAMLYCDAATNKNFNELDQVFL-EDPGQWMDTFVYVEFNPNN--Y-SAVCVSL	339
Db	310	NDVOKAPDGLYKKVNRTEQWYETHQVLPFRPGACITNSAREKINSISLOPDVNLFL	369
Oy	340	GDIIDKVF----RKS-LKG-Y-HSS-LPAPRDKCLPD--QQPIPTETFOVADNHPVA	388
Db	370	KDHFLLMDGOVSRLLLOPAPRYQRAVAVHVGDLHS-TYDVLFELGTGDRLLKATVLSR	428
Oy	389	QRVEPM-GPLKTPLE--HSKYHQKAVAHMAQASHGETHVLYLTDRGTHKKVPEEQ	445
Db	429	VH---I-IEELQIFPGQAPQNLILDSHGGLYASHSHSVYQVPAVANSLSYPT-CGDL	483
Oy	446	EHSEAFNIMEIOFFRRRAAIOITMSDAERKRKLYVSSQWVEVSOVPLDLCEVYGGCHGLM	505
Db	484	ARDPYCAWTGSACRLASLYOPD	505
Oy	506	SRDPYCGMDGRC-I-SYSSSE	525
RESULT	11		

ENTRY S66498 #type complete
TITLE M-sema F protein precursor - mouse
ORANISM #formal_name Mus musculus #common_name house mouse
DATE 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 10-Sep-1997

ACCESSIONS S66498
REFERENCE S66498
#authors Inagaki, S.; Furuyama, T.; Iwahaashi, Y.
#journal FEBS Lett. (1995) 370:269-272
#title Identification of a member of mouse semaphorin family.
#cross-references MIMD:95385809
#accession S66498
#status preliminary
#molecule_type mRNA
#residues 1-834 #label INA

FEATURE #cross-references EMBL:S79463; NID:g1110598; PID:g1110599
1-21
22-834

SUMMARY #length 834 #molecular-weight 92556 #checksum 7189

Query Match 10.5%; Score 522; DB 2; Length 834;
Best Local Similarity 28.5%; Pred. No. 5.15e-83;
Matches 142; Conservative 118; Mismatches 190; Indels 49; Gaps 36;

Db 102 KGKS-NOTECFNIFRLOPNSSHLVYCGTYAFOPKCTYINMLFTLDRAFEDGKGP 160
117 KGSCDLKRDCEYITLLER-RSEGLACGTNARHPS-C-W-NLVNGTV--VPLGEMGYAP 171
Db 161 YDAKAGHTGLVDELXATLN-NFLGTEPVILRMGTHTHSITETYLAEVLPNPHVGA 219
172 FSPDENSLVLEGEDEVYSTIRKGEYNGKIPRRIRIGSESLYSD--TV-MQMPQIKAT 228
Db 220 FVPEVSFGTDGDDKIYFFEFSERAVEYDCYSEQVAVARVAKCKDGGARTIQ-KKMTTF 278
229 IYHQD-QAY---DKITIYFFREDNPKNPAPLNVSRVAKLCRGDGGESSLSVSKWNTF 284
Db 279 LKARLVCSAPDMKVYFNQAKVHTLRGAS--WHNTTFEFGVQARWGDMDLSAVCEYQLDQ 336
285 LKMLVCSAPDATNKNFNLDVFLRDPGSGWMDTRYGVFSNPM---NNSAVCYSLDQ 341
Db 337 IQQVFGSPKKEYSEQAKMARIYDVPSPRPGSCINNMHNDNGYTSLEIPDNTLFIKK 396
342 IDKVF---RT-S-SL-K-G-YHSSLPNRPKGLRDP-QQPIR-TETFOVADHPE-VA- 388
Db 397 HPLMEQVPRGLRPLLVKKNFTNFTYVADRVPGLDGATYVFIETGGGWLKAVSLGP 456
389 QR-VE-PMGF-LKTPLEHSHY-HYQVAVAHMQASHGETFHVLYLTDRGTHIKVVEPGE 444
Db 457 WH--M-VEELOVFOEP-VESELVLSQSKVLFAGSRSQLVOLSLADCTKYRF-CVDCV 510
445 QEHSFAFNIMEIOPFRRAALOTMSLDAERKLYVSSQMEVSVPDLCEVYGGCHGL 504
Db 511 LARDPICAMNVNTSCVATTSGRSGFLVOHVANDTSKMCNOYGIKVRISIPKNTIVS 570
505 MSRDPCYGMND-Q-GRCISITYSER-SVL-QSINPAEPHECPKPKDKA-PLQK-VSLAP 558
Db 571 GTDLVLPCHSSMLAHAW 589
559 NSRIYLSCPWESHATYSW 577

RESULT 12
ENTRY 148745 #type complete
TITLE semaphorin B - mouse
ORANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 28-Feb-1997

ACCESSIONS 148745
REFERENCE 148744
#authors Puschel, A.W.; Adams, R.H.; Betz, H.
#journal Neuron (1995) 14:941-948
#title Murine semaphorin D/collapsin is a member of a diverse gene

#cross-references MIMD:95267431
#accession 148745
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-760 #label RES
#cross-references EMBL:X85991; NID:g854325; PID:g854326

GENETICS
#gene semB
#length 760 #molecular-weight 83458 #checksum 2188

SUMMARY

Query Match 9.7%; Score 480; DB 2; Length 760;
Best Local Similarity 29.2%; Pred. No. 7.12e-74;
Matches 157; Conservative 121; Mismatches 203; Indels 57; Gaps 41;

Db 123 QCFNFRVLSVNTATLVACGTFAPSPACTFIEDLSLLPLIDKVDGKQSLTFT 182
125 DCENYITLLER-RSEGLACGTNARHPS-C-W-NLVNGTVVP-LGE--M--RGYAPFSP-D 175
Db 183 STQAVLVDS-MLYSGT-MNMFSGSEPTLMRLTGLSHPVLTQDIFRLMLH-ADASFVAALPS 239
176 ENSLVLEFGEDEVYSTIRKGEYNGKIPRRIRIGSESLYSDTVMNPQFIKATVHDDA 235
Db 240 -TQVYFFFEETASEDFEEDLYISRNAVCNDVGEKLLQ-KKMTFLKAQLCAQPG 297
236 YDCKIYFFREDNPKNPAPLNVSRVAKLCRGDGGESSLSVSKWNTFLKMLVCSDA 295
Db 298 -QLPFIITHAVLPRDS-P-SVSRIVAVTQONOVGTRSSAVCAFSLTIDERFKGY 354
296 TKKNFRLODVFLLRDPGSGWMDTRYGVFSNPMY---SACVYSGLGIDKVF-R-T- 348
Db 355 KELNETSHMTYRGESEVSPRGSCSMGSPSKALTF-MKD-HFLMDHVVGT-PL---L 408
349 SSL-K-G---YHSLPRPRGKCLPDQPIPTETFOVADHPEVQARVEMGSLKPL 401
Db 409 YKGVETRLAVESARGLDGSSHVMYLGSTGPLHKAVP--QDSS-AYLVEETQLSPD 465
402 FSKHYQKVAVAVHMQASHGETFHVLYLTDRGTHIKVVEPGEQHSFAFNIMEIOPRR 461
Db 466 SEPVRNLQAPRQAGVAFSGSIRVPRANCVES-CVDCVLARDHCAMPDSRLCS 524
462 AAATQMSLDAERKLYVSSQMEVSQVPLDCEVYGGCHGLMSRDYCGND-GR-CI 519
Db 525 LLSGSKPKQDMENGEFVCTRGPMARSPRQSPOLIEVLVPSILELRCPHISA 584
520 SYSSERSVLSQINAEPRKEP-NP--K-PDK-AP--LQKYSLA-PSRIYLSCPWESR 571
Db 585 LASYHMSHGRANISBASATVYNGSL-LLLPDQGVGLY-Q-CVATENGYSYPPVSYW 638
572 HATYSWRH-KENVEGSCERGHQSPCILFIENLTAGQYGHFCEAQEGSY-FREAQHW 627

RESULT 13
ENTRY E42521 #type complete
TITLE A39R protein - vaccinia virus (strain Copenhagen)
ORANISM #formal_name vaccinia virus
#note host Homo sapiens (man)
DATE 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 08-Apr-1994

ACCESSIONS E42521
REFERENCE A33172
#authors Johnson, G.P.
#submission submitted to Genbank, June 1990
#accession E42521
#status preliminary
#molecule_type DNA
#residues 1-403 #label JOH

SUMMARY #length 403 #molecular-weight 45741 #checksum 8167

Query Match 9.5%; Score 470; DB 2; Length 403;
Best Local Similarity 33.5%; Pred. No. 1.04e-71;
Matches 86; Conservative 59; Mismatches 95; Indels 17; Gaps 16;

```
Db 77 LVCGTNGNPKCWK-IDGSDPKHGRGYAPYONSKVTIISYN-ECVLSIDINISK-EG-I 132
| 1111 : 11 : 11 : 1111 : : : : : : : : : :
141 LAGTANARHSCNVLVNGTVPLGEMRGYAPFSPDENSILVFEDEYVSTIRKOEYNGKI 200
| 11 : 11 : 1111 : : : : : : : : : :
133 KWRREFDPCGYDLTYADNVIIPKDG-LRGAFVDDKGYD-KVYILFTDTIGSKR--I-VK 187
| 11 : 11 : 1111 : : : : : : : : : :
201 PRFRRIIGE-S-ELYTSDTWQNQPIKATIVHODQAYDDKIYFFREDNPKNPEAPLN 258
| 11 : 11 : 1111 : : : : : : : : : :
188 IPIIAOMCLNDEGSPSSLSHRMSTFLKVELEC-DID-GMSY-R-Q-IHSRTIKTD-ND 241
| 11 : 11 : 1111 : : 1111 : : : : : : : :
259 VSRVAOICRDQGGESSLSVSKMNTFLKAMLVCSDAATNENFLQVFLPLPSPSGWRD 318
| 11 : 11 : 1111 : : 1111 : : : : : : : :
242 TIIVFEDSPYSKSLATYMTIKOSFSTKLEGYTKOLPSPASGICLPAGKVVSHTFE 301
| 11 : 11 : 1111 : : 1111 : : 1111 : : 1111 : :
319 TRYGYVSNMWNYSAVCVYSLGIDIKVFRITSLKGYHSLPNRPFGKCLPDQPIPIET 378
| 11 : 11 : 1111 : : 1111 : : 1111 : : 1111 : :
Db 302 EVIEKYNVLDIIRPLS 318
| 11 : 11 : 1111 : : 1111 : : 1111 : : 1111 : :
QY 379 QVADRHPEVAORVEPMG 395
| 11 : 11 : 1111 : : 1111 : : 1111 : : 1111 : :

RESULT 14
ENTRY S29921 #type complete
TITLE hypothetical protein 15 - vaccinia virus
ORGANISM #formal_name vaccinia virus
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
09-Sep-1997

ACCESSIONS S29921
REFERENCE S29907
#authors Amegadzie, B.Y.
#submission Submitted to the EMBL Data Library, January 1991
#accession S29921
#status Preliminary
#molecule_type DNA
#residues 1-441 #label AME
#cross-references EMBL:X57318; NID:662239; PID:662254
SUMMARY #length 441 #molecular_weight 50185 #checksum 6034

Query Match 9.3%; Score 463; DB 2; Length 441;
Best Local Similarity 33.9%; Pred. No. 3.37e-70;
Matches 87; Conservative 57; Mismatches 96; Indels 17; Gaps 16;

Db 115 LVCGTNGNPKCWK-IDGSDPKHGRGYAPYONSKVTIISYNGC-VLSDINISK-EG-I 170
| 1111 : 11 : 11 : 1111 : : : : : : : : : :
141 LAGTANARHSCNVLVNGTVPLGEMRGYAPFSPDENSILVFEDEYVSTIRKOEYNGKI 200
| 11 : 11 : 1111 : : : : : : : : : :
171 KWRREFDPCGYDLTYADNVIIPKDG-LRGAFVDDKGYD-KVYILFTDTIGSKR--I-VK 225
| 11 : 11 : 1111 : : : : : : : : : :
201 PRFRRIIGE-S-ELYTSDTWQNQPIKATIVHODQAYDDKIYFFREDNPKNPEAPLN 258
| 11 : 11 : 1111 : : : : : : : : : :
226 IPIIAOMCLNDEGSPSSLSHRMSTFLKVELEC-DID-GMSY-R-Q-IHSRTIKTD-ND 279
| 11 : 11 : 1111 : : 1111 : : : : : : : :
259 VSRVAOICRDQGGESSLSVSKMNTFLKAMLVCSDAATNENFLQVFLPLPSPSGWRD 318
| 11 : 11 : 1111 : : 1111 : : : : : : : :
280 TIIVFEDSPYSKSLATYMTIKOSFSTKLEGYTKOLPSPASGICLPAGKVVSHTFE 339
| 11 : 11 : 1111 : : 1111 : : 1111 : : 1111 : :
319 TRYGYVSNMWNYSAVCVYSLGIDIKVFRITSLKGYHSLPNRPFGKCLPDQPIPIET 378
| 11 : 11 : 1111 : : 1111 : : 1111 : : 1111 : :
Db 340 EVIEKYNVLDIIRPLS 356
| 11 : 11 : 1111 : : 1111 : : 1111 : : 1111 : :
QY 379 QVADRHPEVAORVEPMG 395
| 11 : 11 : 1111 : : 1111 : : 1111 : : 1111 : :

RESULT 15
ENTRY JH0798 #type complete
TITLE fasciclin IV precursor - American bird grasshopper
ORGANISM #formal_name Schistocerca americana #common_name American
bird grasshopper
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
09-Sep-1997

ACCESSIONS JH0798
REFERENCE JH0798
#authors Kolodkin, A.; Matthes, D.J.; O'Connor, T.P.; Patel, N.H.;
```

```
Admon, A.; Bentley, D.; Goodman, C.S.
Neuron (1992) 9:831-845
fasciclin IV: Sequence, expression, and function during
growth cone guidance in the grasshopper embryo.

#journal Admon, A.; Bentley, D.; Goodman, C.S.
#title fasciclin IV: Sequence, expression, and function during
#accession JH0798
#molecule_type mRNA
#residues 1-730 #label KOL
#cross-references GB:100709; NID:6160844; PID:6160845
#experimental_source embryo
#comment This protein plays a role in growth cone guidance in the developing
central nervous system.
KEYWORDS glycoprotein; transmembrane protein
FEATURE
1-72
23-730 #domain signal sequence #status predicted #label SIG\
23-627 #domain fasciclin IV #status predicted #label MAR\
628-652 #domain extracellular #status predicted #label EXT\
653-730 #domain transmembrane #status predicted #label TM\
44,71,163,267,360, #domain intracellular #status predicted #label INT\
539 #binding_site carbohydrate (Asn) (covalent) #status
predicted

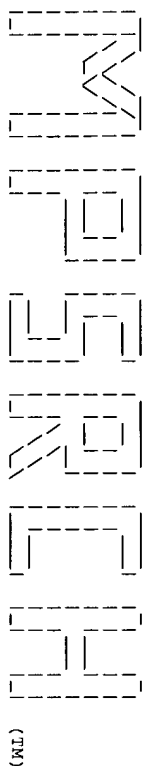
SUMMARY #length 730 #molecular_weight 81214 #checksum 5881

Query Match 9.3%; Score 461; DB 2; Length 730;
Best Local Similarity 28.0%; Pred. No. 9.11e-70;
Matches 143; Conservative 130; Mismatches 182; Indels 56; Gaps 46;

Db 32 QGGEERYQRLNGESHKDFLEKDNHSLGARNIVINISRLDTLEFEQIEVHSSG 91
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
61 HVGQDRVD-FGQTEPHTVLFH--EPGSSVYWGGRKRVYLFDEPEKRN-ASVRTV-NI-G 114
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
92 AHRELVIYLGKSEDDQONTYRLVAKIDDDRVILGCTNAYKPLCRHNAIKDGVVEKEYE 151
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
115 STKGSC-L-DKR--DCENITTLDERSEG-LLCGNNARHPSQW--LVNGT-VVPLG-E 165
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
152 GRGLCFDPDPDHNSAIYSEGLYSATV-ADFSQTD-LI-YRG-P-LTERSDL-R-QLN 204
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
166 MRGYAFSPDENSILVFEDEYVSTIRKOEYNGKIYFFREDNPKNPEAPLN-SDTWQNQPI 225
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
205 APNFEVN-TMEYNDPIFEFFRETAIEYINGKAIY-SRVARVCHDKGPHQFG-DKWTSE 261
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
226 KATIVHODQAYDDKIYFFREDNPD-KNPEAPLVNRSVQOLCRDQGGESSLSVSKMNT 284
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
262 LKSLRNCVYGDYPFYFNEIIOSTDIIEGYGQ-VEKLIYGVFTVPVNSIGSACAFS 320
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
285 LKAMLVCSDAATNK-NENRLQVDF-LLPDP-SQMDRTRYGVFSNPMW-Y--SAVCVYS 338
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
321 MKSILSFDPDFKEQETMNSMNLAVPSLVKPEPQCVDNSRTLPDVSYNFKSHITLMD 380
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
339 L-G--D-IDKVFRT-SSLKG-YHS--SL--PNRPKCLPDQPIPIETFGQVADRHPEVA 388
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
381 EAVPAEFT-PIIIRISLOYRFTKIIVDOQVTPDGKAVDLFISGDDGKVIALNSASF 439
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
389 QREVPMPGLTKPLFH-S-KHYOKVAV-HRMQASHETFEHLVLTDRGIHRYVPEPG-- 443
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
440 DSSDIDSVYIEELQVLPQVYKLVYVMDSDSLVVSDEDLAIKILHRGSDKIT 499
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
444 ECHSFAFNIME-IQPRRAAIIQIM---SLDERKRLVYSSQWEVSQVPLDCEVYG-G 498
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
500 NCRECVSLDPPYCAMDNVELCTAVGSPDMS 530
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
499 GCHGLMSRDPYCGMDQ-G-RCISISYSSERS 527
| 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :

Search completed: Tue Aug 3 16:28:41 1999
Job time : 110 secs.
```

This Page Blank (uspto)



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Aug 3 16:32:43 1999; Maspar time 8.46 Seconds
799,386 Million cell updates/sec
Tabular output not generated.

Title: >US-09-240-410-2
Description: (1-666) from US09240410.pep
Perfect Score: 4968
Sequence: 1 MTPPPGGAAPSAPRARVPG.....LAASIMGLVPLTIGLGVH 666

Scoring table:
PAM 150
Gap 11

Searched: 106580 seqs, 10152877 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCR9_COMB 4:backfiles1

Statistics: Mean 35.201; Variance 163.168; scale 0.216

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the predicted of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	661	13.3	771	3	PCT-US94-1	Sequence 54, Applicati	9,28e-47
2	661	13.3	771	2	US-08-835-	Sequence 54, Applicati	9,28e-47
3	661	13.3	771	1	US-08-121-	Sequence 54, Applicati	9,28e-47
4	620	12.5	477	1	US-08-136-	Sequence 2, Applicatio	4,24e-43
5	463	9.3	441	3	PCT-US94-1	Sequence 56, Applicati	3,22e-29
6	463	9.3	441	2	US-08-835-	Sequence 56, Applicati	3,22e-29
7	463	9.3	441	1	US-08-121-	Sequence 56, Applicati	3,22e-29
8	459	9.2	730	1	US-08-835-	Sequence 58, Applicati	7,22e-29
9	459	9.2	730	1	US-08-121-	Sequence 58, Applicati	7,22e-29
10	459	9.2	730	1	PCT-US94-1	Sequence 58, Applicati	7,22e-29
11	378	7.6	712	1	US-08-121-	Sequence 64, Applicati	7,91e-22
12	378	7.6	712	2	US-08-835-	Sequence 64, Applicati	7,91e-22
13	378	7.6	712	2	PCT-US94-1	Sequence 64, Applicati	7,91e-22
14	367	7.4	650	1	US-08-121-	Sequence 60, Applicati	7,00e-21
15	367	7.4	650	3	PCT-US94-1	Sequence 60, Applicati	7,00e-21
16	367	7.4	650	2	US-08-835-	Sequence 60, Applicati	7,00e-21
17	362	7.3	724	1	PCT-US94-1	Sequence 62, Applicati	1,88e-20
18	362	7.3	724	1	US-08-121-	Sequence 62, Applicati	1,88e-20
19	362	7.3	724	2	US-08-835-	Sequence 62, Applicati	1,88e-20
20	157	3.2	122	2	US-08-835-	Sequence 66, Applicati	1,39e-03
21	157	3.2	122	1	US-08-121-	Sequence 66, Applicati	1,39e-03
22	157	3.2	122	3	PCT-US94-1	Sequence 66, Applicati	1,39e-03
23	115	2.3	422	3	PCT-US95-0	Sequence 170, Applicat	1,67e+00

24	115	2.3	422	2	US-08-469-	Sequence 170, Applicat	1,67e+00
25	115	2.3	422	2	US-08-339-	Sequence 3, Applicatio	1,67e+00
26	115	2.3	422	3	PCT-US94-0	Sequence 166, Applicat	1,67e+00
27	115	2.3	422	2	US-08-428-	Sequence 3, Applicatio	1,67e+00
28	115	2.3	422	2	US-08-428-	Sequence 3, Applicatio	1,67e+00
29	115	2.3	422	2	US-08-734-	Sequence 170, Applicat	1,67e+00
30	115	2.3	422	1	US-08-036-	Sequence 170, Applicat	1,67e+00
31	115	2.3	422	1	US-08-469-	Sequence 170, Applicat	1,67e+00
32	115	2.3	422	1	US-08-249-	Sequence 170, Applicat	1,67e+00
33	115	2.3	422	1	US-08-428-	Sequence 3, Applicatio	1,67e+00
34	115	2.3	422	1	PCT-US94-0	Sequence 185, Applicat	1,67e+00
35	115	2.3	422	2	US-08-469-	Sequence 170, Applicat	1,67e+00
36	107	2.2	533	1	US-08-445-	Sequence 10, Applicati	5,97e+00
37	107	2.2	533	2	US-08-484-	Sequence 13, Applicati	5,97e+00
38	107	2.2	533	1	US-08-484-	Sequence 13, Applicati	5,97e+00
39	100	2.0	780	1	US-08-232-	Sequence 14, Applicati	1,77e+01
40	100	2.0	780	2	US-08-786-	Sequence 14, Applicati	1,77e+01
41	100	2.0	1311	2	US-08-340-	Sequence 5, Applicatio	1,77e+01
42	94	1.9	146	1	US-07-795-	Sequence 26, Applicati	4,40e+01
43	94	1.9	146	1	US-08-457-	Sequence 26, Applicati	4,40e+01
44	92	1.9	321	3	PCT-US94-1	Sequence 26, Applicati	5,93e+01
45	92	1.9	857	1	US-07-717-	Sequence 3, Applicatio	5,93e+01

ALIGNMENTS

RESULT 1
ID PCT-US94-10151A-54 STANDARD: PRT: 771 AA.
XX xxxxxx
AC
XX
DT
XX
DE Sequence 54, Application PC/TUS9410151A
CC GENERAL INFORMATION:
CC APPLICANT: The Regents of the University of California
CC TITLE OF INVENTION: The Semaphorin Gene Family
CC NUMBER OF SEQUENCES: 66
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: FLEHR HOBBACH TEST ALBRITTON & HERBERT
CC STREET: 4 Embarcadero Center, Suite 3400
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94111-4187
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/10151A
CC FILING DATE: 13-SEP-1994
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Osman, Richard A.
CC REGISTRATION NUMBER: 36,627
CC REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 781-1989
CC TELEFAX: (415) 398-3249
CC TELEX: 910 277299 FHT UR
CC INFORMATION FOR SEQ ID NO: 54:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 771 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 771 AA; 88889 MW; 2976250 CN;

Query Match 13.3%; Score 661; DB 3; Length 771;

Best Local Similarity 30.3%; Pred. No. 9,286-47;
Matches 151; Conservative 126; Mismatches 174; Indels 47; Gaps 37;

D 166 RGSFDPKLLTASLLIDGELYSCTA-ADFMGRDFAIFRTLGHHNPIRTEQHSRWLNDP 224
Q 167 RGAAPSPDENSLVLEGEDEVSTIRKOEYNGK-IPFRRIKGESELYTS--DTV-MQNP 222
D 225 KFIASHLISEDNEDDKYIFFERENALIDGESHGKATHRIGQICKNDGGHSL-VNWK 283
Q 223 QFIKATIVHQ-DOAYDKKITYFFREDNPDKNPEAPLNVSRVAQLCRDGGESSLSVSKW 281
D 284 TTFKLARLISVPGPNIGDTHFDELQDVFLM-NFKDP-KNPVYGVTTSSNIFKSAVC 341
Q 282 NTFKLAMLYCS--DA-ATKNFNRLODVFLLPDPSGQMDTRVYGVFSNPMW-Y--SAVC 335
D 342 MYSMSVRRVFLGPIYAHNRGPNQWVYQGRVYPRPGTCSKTFGGFSDTKDLPDVIT 401
Q 336 VYSLGIDIDKVFRTS-SLK-G--YH--S--S-LPNPRGKCLPDQ-QPI-PTETF--QV-A 381
D 402 FARSHPMYNPVPMNNRPIVITKIDVNYQFTQIYDVRVAEDGQ-YDVMFICTDVGTVLK 460
Q 382 -DR-HPEVAQRVEPMGLKPL-FHSKTHYQKVAVHRMQASHGETFHVLYLTTRDGTIRK 438
D 461 VVSIKRETWYDLEEVLLLEMTVREPTAISAMELSTKQOQLYIGSTAGVAQLPLHRCDIY 520
Q 439 VVE-PGEDEHSAFNMIE-IQPFRRAAIQTMSLDAERKLYVSSQWESQVPLDCEVY 496
D 521 KACAECCCLARDPYCAMDGSACSRFPYAKRRTRODIRNGDPLTHCSDLHNDHGHSP 580
Q 497 GGGCHGCLMSRDPYCGMDGRCISYSSSE-RSVL-QSINPAEPHKECPNPKPKD----AP 550
D 581 EERLIYGVENSSTFLECSPKSQRALVYWOFORNEERKEIRVDHIIIRTDGILLRSIQ 640
Q 551 LQKVSIA-PNSRYLYSC-PMESRHATY-SW-RHKEVY-EQSCPEHQ-SPNCILFIENLT 604
D 641 QKDSGNVLCHAVHEGFIQ 658
Q 605 AQOYGHYFCEAOEGSYFR 622

RESULT 2
ID US-08-835-268-54 STANDARD: PRT: 771 AA.
AC xxxxxx
XX
DT
XX
XX
XX
Sequence 54, Application US/08835268
CC Sequence 54, Application US/08835268
CC Patent No. 5807826
CC GENERAL INFORMATION:
CC APPLICANT: Goodman, Corey S.
CC APPLICANT: Kolodkin, Alex L.
CC APPLICANT: Matthes, David
CC APPLICANT: Bentley, David R.
CC APPLICANT: O'Connor, Timothy
CC TITLE OF INVENTION: The Semaphorin Gene Family
CC NUMBER OF SEQUENCES: 100
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
CC STREET: 268 Bush Street, Suite 3200
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/835,268

CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/121,713
CC FILING DATE: 13-SEP-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Osman, Richard A.
CC REGISTRATION NUMBER: 36,627
CC REFERENCE/DOCKET NUMBER: B94-002-1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415)343-4341
CC TELEFAX: (415) 343-4342
CC
CC INFORMATION FOR SEQ ID NO: 54:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 771 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 771 AA; 88889 MW; 2976250 CN;

Query Match 13.3%; Score 661; Db 2; Length 771;
Best Local Similarity 30.3%; Pred. No. 9,286-47;
Matches 151; Conservative 126; Mismatches 174; Indels 47; Gaps 37;

D 166 RGSFDPKLLTASLLIDGELYSCTA-ADFMGRDFAIFRTLGHHNPIRTEQHSRWLNDP 224
Q 167 RGAAPSPDENSLVLEGEDEVSTIRKOEYNGK-IPFRRIKGESELYTS--DTV-MQNP 222
D 225 KFIASHLISEDNEDDKYIFFERENALIDGESHGKATHRIGQICKNDGGHSL-VNWK 283
Q 223 QFIKATIVHQ-DOAYDKKITYFFREDNPDKNPEAPLNVSRVAQLCRDGGESSLSVSKW 281
D 284 TTFKLARLISVPGPNIGDTHFDELQDVFLM-NFKDP-KNPVYGVTTSSNIFKSAVC 341
Q 282 NTFKLAMLYCS--DA-ATKNFNRLODVFLLPDPSGQMDTRVYGVFSNPMW-Y--SAVC 335
D 342 MYSMSVRRVFLGPIYAHNRGPNQWVYQGRVYPRPGTCSKTFGGFSDTKDLPDVIT 401
Q 336 VYSLGIDIDKVFRTS-SLK-G--YH--S--S-LPNPRGKCLPDQ-QPI-PTETF--QV-A 381
D 402 FARSHPMYNPVPMNNRPIVITKIDVNYQFTQIYDVRVAEDGQ-YDVMFICTDVGTVLK 460
Q 382 -DR-HPEVAQRVEPMGLKPL-FHSKTHYQKVAVHRMQASHGETFHVLYLTTRDGTIRK 438
D 461 VVSIKRETWYDLEEVLLLEMTVREPTAISAMELSTKQOQLYIGSTAGVAQLPLHRCDIY 520
Q 439 VVE-PGEDEHSAFNMIE-IQPFRRAAIQTMSLDAERKLYVSSQWESQVPLDCEVY 496
D 521 KACAECCCLARDPYCAMDGSACSRFPYAKRRTRODIRNGDPLTHCSDLHNDHGHSP 580
Q 497 GGGCHGCLMSRDPYCGMDGRCISYSSSE-RSVL-QSINPAEPHKECPNPKPKD----AP 550
D 581 EERLIYGVENSSTFLECSPKSQRALVYWOFORNEERKEIRVDHIIIRTDGILLRSIQ 640
Q 551 LQKVSIA-PNSRYLYSC-PMESRHATY-SW-RHKEVY-EQSCPEHQ-SPNCILFIENLT 604
D 641 QKDSGNVLCHAVHEGFIQ 658
Q 605 AQOYGHYFCEAOEGSYFR 622

RESULT 3
ID US-08-121-713D-54 STANDARD: PRT: 771 AA.
AC xxxxxx
XX
DT
XX
XX
XX
Sequence 54, Application US/08121713D
CC Sequence 54, Application US/08121713D
CC Patent No. 5639856

[illegible]

OY	551	LQKVSIA-PNSRYIYSYC-PMESHATY-SW-RKEKNV-EQSCPEHQ-SPNCIFIEMLT	604
DB	641	OKDSGNYLCHAVEHGFQ	658
OY	605	AQQGYHFCEAOEGSYFR	622
RESULT	4		
ID	US-08-136-922-2	STANDARD:	PRT: 477 AA.
AC	xxxxxx		
XX			
XX			
DE	Sequence 2, Application US/08136922		
CC	Sequence 2, Application US/08136922		
CC	Patent No. 5416197		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Kaper, Jonathan A.		
CC	APPLICANT: Luo, Yuling		
CC	TITLE OF INVENTION: Compositions Which Regulate Neural		
CC	TITLE OF INVENTION: Regeneration and Methods of Making and Using the Same		
CC	NUMBER OF SEQUENCES: 2		
CC	CORRESPONDENCE ADDRESSES:		
CC	ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &		
CC	ADDRESSEE: No. 5416197is		
CC	STREET: One Liberty Place		
CC	CITY: Philadelphia		
CC	STATE: PA		
CC	COUNTRY: USA		
CC	ZIP: 19103		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: Patentin Release #1.0, Version #1.25		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/136,922		
CC	FILING DATE:		
CC	CLASSIFICATION: 435		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Deluca, Mark		
CC	REGISTRATION NUMBER: 33, 229		
CC	REFERENCE/DOCKET NUMBER: UPN-1428		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 215-568-3100		
CC	TELEFAX: 215-568-3439		
CC	INFORMATION FOR SEQ ID NO: 2:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 477 amino acids		
CC	TYPE: amino acid		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: Protein		
CC	SEQUENCE 477 AA; 54647 MW; 1161988 CN;		
SO			
Query Match	12.5%; Score 620; DB 1; Length 477;		
Best Local Similarity	31.2%; Pred. No. 4,24e-43;		
Matches	134; Conservative 110; Mismatches 145; Indels 41; Gaps 32;		
DB	1 LEHDNEDDKYFFEFRENAIDGEHSKATHARIGQICNDFGGHSL-VNKWTTFLLARL	59	
OY	230 VHQDADYDKIYYFREDNPDKNEAPLANVSRAQLCRDGOGSGSSLSVSKWNFLXAML	289	
DB	60 TCSVEGPNSIDIHFDLOVFLM-NFKDP-KNVVYGVFTTSNIIFGSAVCMTSMGDVR	117	
OY	290 VCS--DA-ATNNFNELQVFLLPDPSGOWRDRYGVFSNPNW-Y--SACVYSLSDID	343	
DB	118 RYFLGLGARHPNPNYMWVYOGRVPRPOTCPSKTGFGDFDTPDLDDPVITTFARSHPA	177	
OY	344 KYFRIS-STLK-G-YH-S-S-LPNRPBROCKCLPDQ-QPI-PIETTF--QV-A-DR-HPEV	387	

[illegible]

QY	141	LACGTNRHRSCHMLVNGTIVPGLGEMRGVAFPSFSPDENSELVLFEGDEYVSTIRKQEYNGKI	200
Db	171	KRMREDPGCGDLYTADNVIPKDG-LRGAFVDKDGYYD-KVYILFDITGSKR--I-VK	225
QY	201	PRFRRIIGE-S-ELYETSDTVMQNPQFIKATIVHQDAIYDKIKYIFREDNPDPNPEAPLN	258
Db	226	IPFIAMCUNLDEGSSLSHRMSTFLKVELE- DID-GRSY-R-O-IHSRTIKTD-ND	279
QY	259	VSNVAOLCRDDGGESSLSVSKWNTFLKMLWCSDAATNKNENRLODFVLLPDPSSQWMD	318
Db	319	TRYGVFSNPMWNSAACYVSLGDIDKVEFTSSLKGYHSSLPNRPCKLPDQPIETET	378
QY	340	EVLKINVLDDIIRKPLS	356
QY	379	QVADRHPEVAQRVEPMWG	395
RESULT	6		
ID	US-08-835-268-56	STANDARD;	PRT; 441 AA.
AC	xxxxxx		
XX			
DT			
XX			
DE			
XX			
Sequence 56, Application US/08835268			
CC	Sequence 56, Application US/08835268		
CC	Patent No. 5807826		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Goodman, Corey S.		
CC	APPLICANT: Koloedkin, Alex L.		
CC	APPLICANT: Matches, David		
CC	APPLICANT: Bentley, David R.		
CC	APPLICANT: O'Connor, Timothy		
CC	TITLE OF INVENTION: The Semaphorin Gene Family		
CC	NUMBER OF SEQUENCES: 100		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP		
CC	STREET: 268 Bush Street, Suite 3200		
CC	CITY: San Francisco		
CC	STATE: CA		
CC	COUNTRY: USA		
CC	ZIP: 94104		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/835,268		
CC	FILING DATE:		
CC	CLASSIFICATION:		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/121,713		
CC	FILING DATE: 13-SEP-1993		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Osman, Richard A.		
CC	REGISTRATION NUMBER: 36,627		
CC	REFERENCE/DOCKET NUMBER: B94-002-1		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: (415)343-4341		
CC	TELEFAX: (415) 343-4342		
CC	TELEX:		
CC	INFORMATION FOR SEQ ID NO: 56:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 441 amino acids		
CC	TYPE: amino acid		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: protein		
CC	SEQUENCE 441 AA; 50185 MW; 1082834 CN;		
QY	SEQUENCE		


```

XX XX Sequence 58, Application PC/TUS9410151A
XX DE
CC CC Sequence 58, Application PC/TUS9410151A
CC CC GENERAL INFORMATION:
CC CC APPLICANT: The Regents of the University of California
CC CC TITLE OF INVENTION: The Semaphorin Gene Family
CC CC NUMBER OF SEQUENCES: 66
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: FLEHR HOHBACH TEST ALBRITTON & HERBERT
CC CC STREET: 4 Embarcadero Center, Suite 3400
CC CC City: San Francisco
CC CC STATE: CA
CC CC COUNTRY: USA
CC CC ZIP: 94111-4187
CC CC COMPUTER READABLE FORM:
CC CC MEDIUM TYPE: Floppy disk
CC CC COMPUTER: IBM PC compatible
CC CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CC CURRENT APPLICATION DATA:
CC CC APPLICATION NUMBER: PCT/US94/10151A
CC CC FILING DATE: 13-SEP-1994
CC CC CLASSIFICATION:
CC CC ATTORNEY/AGENT INFORMATION:
CC CC NAME: Osman, Richard A.
CC CC REGISTRATION NUMBER: 36,627
CC CC REFERENCE/DOCKET NUMBER: EP-58750-PC/RAO
CC CC TELECOMMUNICATION INFORMATION:
CC CC TELEPHONE: (415) 781-1989
CC CC TELEFAX: (415) 398-3249
CC CC TELEX: 910 277299 FHT UR
CC CC INFORMATION FOR SEQ ID NO: 58:
CC CC SEQUENCE CHARACTERISTICS:
CC CC LENGTH: 730 amino acids
CC CC TYPE: amino acid
CC CC TOPOLOGY: linear
CC CC MOLECULE TYPE: protein
SQ SEQUENCE 730 AA; 81124 MW; 2758267 CN;
Query Match 9.2%; Score 459; DB 3; Length 730;
Best Local Similarity 28.0%; Pred. No. 7,22e-29;
Matches 143; Conservative 130; Mismatches 182; Indels 56; Gaps 46
Db 32 QGEEYVORFGNESHKHFKLEEDHNSLLVGARNIYYNLSLDLTFEETFORIEWHSSG 91
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 HVGODRVD-FQOTEPRHYLFH--EPGSSVWVGGRGKYILFDPEEGKN-ASVRIV-NI-G 114
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 AHRELCLYLGKSEDDCCQYIRVLAKIDDRYLIGGTNAKPLCRHVALKDGDVVEKEYE 151
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 STKGSC-L-DKR--DCENYFTLLERBSRG-LLAGGTNARHPSCWN--LVNGT-VVPIG-E 165
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 152 GRGLCPFPDPDINSRIAYISEGQLYSATV-ADSSGTD-P-LI-YRG-P-LITEISDL-K-QLN 204
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 166 MRGAAPSPDENSSIVLFEFGDEVYSTIRKOEFNGKIPRRIRGSSELYTSDTVQNQOFI 225
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 APNVPAF-TMENVDFEFFEETFRETAVEYINCGKAIF-SRYARYCKRHDKG-PHOGDRTTSF 261
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 KATIYHQDAVDXKIITYFFREDND-KNPFLPWLVSRAQCLCRDGOGESSLSISKNTF 284
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 262 LKSRLNCSPGDPYPFYFNELIOSTSDIIIEGNYGGO-VEKLIVGETTPVNSIGSGAVCAFS 320
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 LKAMLVCSDAATNK-NENRLQDV-F-LPDP-SGMWRDTRVYGVPSNPMN-Y--SAVCVYS 338
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 321 MKSLIESDFGFKEDETNSNMVLAVPSLKAVEPPRPQCVDNSRTLPLPDVSVNFVASHLMD 380
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 339 L-G-D-IDKFERT--SSLKG-YHS--SL--PNPFRPKLPDQOPIPTEIFQVADRHEVA 388
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 381 EAVPAFF-RPIILISIQYRFRTIAVOOVQRTPDGKAYDVLFTGTDGKVIKALNSASF 439
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 389 ORVDEMGCLKPLRHL-S-KIHOKYAA-HRMQASHGEFFHVLIITLRDGIITHKVVVEG-- 443

```

[illegible]

```

Db      166 SYNG-OLEFSATV-ADSSGGDPLIYRPORTEL-SDLKOLANAFNVS--V---ARGDYI 216
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      181 LEFGDEVYITIRKOEYNGATIPRRRRRGESSELYTSDTWMQNPQFIKATIVHQDQAYDKI 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      217 FFFRYETAWEYNMGKVIV-SRYARYCKDKDGG-PHOSRDRWTSPLKRLNCISIGEYF 274
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      241 YFFERBDNDP-KNPEAPLNVSRVAOLCNRDQGESLSVSKMNTFLKAMLYCSDAATNK- 298
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      275 YPEIOTSDIYEGRANSDSKKIITGITTVPVNAIGGSAICAYOMADILVFEGSFKHQ 334
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      299 NFNRIODVFLDPDSCGOMBDR--VYGFVSNFPMN-Y--SACVYSLGIDKVF----R-T 348
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      335 ETNKNMWLEPVQNLVPEPRPGOCVDSRILDPKNNNFIKTISLMD-VPALFG-KPLVLR 392
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      349 SSIKCG-Y-H-S-SL-PNRPFGKCLPDQOPIPTETIFQVADRHPVGAORPEKPLKPTLFH 403
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      393 VSLQYRFAITAYDPOVKITINNLYDLVLTIGTDDGKVLKAVNIPRHKALLYKRYRFSVH 452
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      404 -S-KYHQKVAVH-RMQASHGETFHVLLITDRGIRHKVVE-PBQESHFAFNIMETQF 459
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      453 PHGAPYKOLKIAPGCGKVVVVGKDEIRLANINHCAS-KTRCKDCVELDHPHCAMDAKONT 511
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      460 RRAAAIQTMSLAEERKLYVSSQWVPLDCEVYGCGCHGLMSRDPYCGMD--QGR 517
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      512 CVSIDVTMSYRLLIDVYVRGDNDK-CWSPQTDKKTIVK 548
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      518 CISTIS-SE-RSVLOSINPAEPHKECPNPKPDKAPLOK 553
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
ID      US-08-835-268-64          STANDARD:          PRT:          712 AA.
XX      xxxxxx

DE      Sequence 64, Application US/08835268

CC      Sequence 64, Application US/08835268
CC      Patent No. 5807826
CC      GENERAL INFORMATION:
CC      APPLICANT: Goodman, Corey S.
CC      APPLICANT: Kolodkin, Alex L.
CC      APPLICANT: Mathes, David
CC      APPLICANT: Bentley, David R.
CC      APPLICANT: O'Connor, Timothy
CC      TITLE OF INVENTION: The Semaphorin Gene Family
CC      NUMBER OF SEQUENCES: 100
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
CC      STREET: 268 Bush Street, Suite 3200
CC      CITY: San Francisco
CC      STATE: CA
CC      COUNTRY: USA
CC      ZIP: 94104
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patentin Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/835_268
CC      FILING DATE:
CC      CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/121_713
CC      FILING DATE: 13-SEP-1993
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Osman, Richard A.
CC      REGISTRATION NUMBER: 36,627
CC      REFERENCE/DOCKET NUMBER: B94-002-1
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (415)343-4341

```

CC	TELEFAX:	(415) 343-4342
CC	TELEX:	
CC	INFORMATION FOR SEQ ID NO:	64:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	712 amino acids
CC	TYPE:	amino acid
CC	TOPOLOGY:	linear
CC	MOLECULE TYPE:	protein
CC	SEQUENCE	712 AA; 79751 MW; 2715572 CN;
Query Match	7.6%;	Score 378; DB 2; Length 712;
Best Local Similarity	27.0%;	Pred. No. 7,91e-22;
Matches	140; Conservative	120; Mismatches 194; Indels 56; Gaps 43.
Db	HLTVLNO-DETSLVGNRRYVNLSIFDLSEKGGRIDMPS-SDAHGOLCLLKGTDDC	105
Oy	I : : : : I : : : : I : : : : I : : : : I : : : : I : : : : I : : : : I : : : :	
75	HVLFEHPSSSWVGGRGVY-L--FDPECKNASVTAVNGSRKS-C-L-DKRD--C	126
Db	QNTIRILYSSEEGKLIVCGTNSRKCLTYAFKEKKIYEVEVGIGLCIPNPDHSTSV	165
Oy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
127	ENTITLLERRSG-LLAGCTNARHPCSNLV--NGT-VVPLG-DMRGAPSPDENSL-V	180
Db	SYNG-OLFSAFY-ADESGDDLPIYEPRTEL-SDLQOLNAPFNVS--V----AYGYI	216
Oy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
181	LREGDEVYTIKOENKGIPIFRRIKGESELYISDYVMOPFIKATIVHQDAYDKI	240
Db	FFPYRETAVERNCGKVLY-SRAVARCKDDKG-PHOSRDRWTSFLKARLNCSTPGEPF	274
Oy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
241	YFFREDND-KNPAPLNVRVAQLCRGDGGSLSYSKMTEFLKAMLYCSDAATNK-	298
Db	YDELQISDIYEGKYNDSDKIIYLITTPVNAIGSALICAYOMADILRVGESEKHQ	334
Oy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
299	NNRIADVFLLPDPGWRDTR--VYGFSNPMN-Y--SAVCYSTGLIDIKVF---R-T	348
Db	EFINSNLVPQNLYEPRPGOCYRDSRLILDDKNVNFITKHSIMED-VPLFG-KPLYLR	392
Oy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
349	SSLKG-Y-H-S-SL-PNRPGKCLPDOPILETFTGYADRRPEVAORVEEMGPLKPLFH	403
Db	VSLQRFRTAIVDPVKTIINNQQYLDVLYIGTDGCVLKAVNIPIRHKALLRYKRTSVH	452
Oy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
404	-S-KHYOKVAVH-RMQASHGETFHVLYLTDRGITIKHYE-FEODSHSAFINMELOPF	459
Db	PGADEVKOLKTAPGYGKVVNVGKDEIRIANLNHCAS-KTRCKDCVELQDPHCAMDAKONT	511
Oy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
460	RRAAAIQTMSDIAERKKLYVSSQWEVAGQVPLDCEYVGGCGHGCLMSRDYCGMD--QG	517
Db	CYSIDTIVSYRFLIDOVYRGDNK-CWSPTQDKTYVK	548
Oy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
518	CISTIS-SE-RSVLOSINPAEPHECPNPKPADKAPLOK	553
RESULT	13	
ID	PCr-US94-10151A-64	STANDARD: PRT: 712 AA.
AC	xxxxxx	
DT		
XX		
DE	Sequence 64, Application PC/TUS9410151A	
CC	GENERAL INFORMATION:	
CC	APPLICANT: The Regents of the University of California	
CC	TITLE OF INVENTION: The Semaphorin Gene Family	
CC	NUMBER OF SEQUENCES: 66	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: FIEHR HOHBACH TEST ALBRITTON & HERBERT	
CC	STREET: 4 Elmharcadero Center, Suite 3400	
CC	CITY: San Francisco	
CC	STATE: CA	
CC	COUNTRY: USA	
CC	ZIP: 94111-4187	
CC	COMPUTER READABLE FORM:	

```

CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/US94/10151A
CC      FILING DATE: 13-SEP-1994
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Osman, Richard A.
CC      REGISTRATION NUMBER: 36,627
CC      REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (415) 781-1989
CC      TELEFAX: (415) 398-3249
CC      TELEX: 910 277299 FHT UR
CC      INFORMATION FOR SEO ID NO: 64:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 712 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 712 AA; 79751 MW; 2715572 CN;
SQ

Query Match          7.6%; Score 378; DB 3; Length 712;
Best Local Similarity 27.0%; Pred. No.7,91e-22;
Matches 140; Conservative 128; Mismatches 194; Indels 56; Gaps 43.

Db      48 HFIYINQ-DENSILVGGNRYNLSTIPDLSERKGRIDWPS-SDAHGOLCIKKTDDC 105
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      75 HTVLFHEGSSSVWVGSGKYY-L--EDFPGKNAVPTVIGSTKS-C-L-DKRD--C 126
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      106 QNTRIILYSEPGKLVIGTNSYKPLCTYFAKGGKLYVEKEVIGLCPEYNPHNSTV 165
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      127 ENYITLLERSEEG-LLACGTNARHSPSCWNLV--NGT--VPLIG-EMRGVAPSPDENSL-V 180
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      166 SYNG-QLFSAIV-ADFSGCDPLITREPORITL-SDKOLNAPNTVNS--V---AYGDIY 216
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      181 LEEGDEVYSTIRKOEYNKIPRFRRIGSESLYSDTVMONPOPIKATVHQDAYDCKI 240
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      217 FFFYRETAVEYVNCCKVLY--SRVARVCKDDKGG--PHOSRDRWTSFLKARLNCSPGEYF 274
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      241 YFFREDNPD-KNEPAPLANSRVAQLCRGDGGESSLSVSKNITFLAMLYCSPAATNK- 298
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      275 YFDEIQTSDIVEGRKNSDSDKIITYGILTPVNAIGSALICAYQAMADILVEFGSEKHQ 334
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      299 NENRLQDVELLPDPSGQQRDIR--VYGFSPMN-Y--SAVCYSLGDIDKVF---R-T 348
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      335 ETISNNLVPQNLVPERPGCCVRSKILPDKVNVNFKTHSLMED--VPALFG-KPVLVR 392
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      349 SSLGK-Y-H-S-SL-PNPRPGKCLPDDQPIPTETFOVADRHPVACQHEVEPWLKPLTGL 403
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      393 VSLQRFALIVDDQVKINNOYLIVLYIGDDCKVLKAVNIPRKHAKALLYRKYRFSVH 452
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      404 -S-KTHQKAVH-RMQASHGETHVALYLLTDKGTIRHVE--PEQSHSFAFINMIETQPF 459
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      453 PHGAPVKOLKIAPGYKVVVVGKDEIRLANINHCAS-KTRCKDCVEVLQDPHCAMDAKNTL 511
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      460 RRAALIQMSIDAERKRIYVSSQGEVSGVPLDLCEYVGGGCHGLMSRDPIYCGMD--QGR 517
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      512 CVSIDYVTSYRFLIQDVVRGDNK-CWSPQTDKKTIVK 548
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      518 CISIYS-SE-RSVLQISINPAPHPHKECPNPKDKAPLQK 553
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
ID      US-08-121-713D-60      STANDARD;      PRT;      650 AA.
XX      xxxxxx
XX      DT
XX      DE
Sequence 60, Application US/08121713D

```

```

XX Sequence 60, Application US/08121713D
CC Patent No. 563856
CC GENERAL INFORMATION:
CC APPLICANT: Goodman, Corey S.
CC APPLICANT: Koldkin, Alex L.
CC APPLICANT: Matthes, David
CC APPLICANT: Bentley, David R.
CC APPLICANT: O'Connor, Timothy
CC TITLE OF INVENTION: The Semaphorin Gene Family
CC NUMBER OF SEQUENCES: 100
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
CC STREET: 268 Bush Street, Suite 3200
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/121,713D
CC FILING DATE: 13-SEP-1993
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Osman, Richard A.
CC REGISTRATION NUMBER: 36,627
CC REFERENCE/DOCKET NUMBER: B94-002-1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415)343-4341
CC TELEFAX: (415) 343-4342
CC TELEX:
CC INFORMATION FOR SEQ. ID NO.: 60:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 650 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 650 AA; 72940 MW; 2193067 CN;
QY Query Match 7.4%; Score 367; DB 1; Length 650;
QY Best Local Similarity 28.4%; Pred. No. 7,00e-21;
QY Matches 124; Conservative 104; Mismatches 158; Indels 51; Gaps 38;
Db 3 DCQVIRIMVPSGRLFCVGTNSFRMCMNTYIISDSNTYLLEATKNGQACPIDPDRNST 62
QY 125 DCENVITTLERRSG-LIACGTNARHSCWN-LVNGTVVPLGEWR-GYA--PSPDENS 179
Db 63 SVLADNELYSGTV-ADFSGSDPIYRPLDTEOYDSLS-LNAPNFV-SSFT-OGD-F--- 114
QY 180 VLESGDEYVITIRKQETNGKILPRRRIRGSESLYTSPTWQNPQFIKATTVHODQAYDK 239
Db 115 VYFFRETAVFIFNGKRAIY-SRYAVCYKWDKGGPARFR-NRMTSFLKSLNCSIPGDYP 172
QY 240 IYFFREDNPD-KNPEAPLNVSRVAQICRGDGGSESLSVSKNMTFLKALVLCSDATNK 298
Db 173 FYFNEIQSASLVVGQIGSMSSKILIVENTPNSIIPGSAVCAFALODIADTEGQFKQ 232
QY 299 -NFRLODFV-LLEPDSGQWDRFVYVFSNPNV-Y--SAVCYSLGDI-DKY---FR-- 347
Db 223 TGINSNMPLPVNNAVPRPRGSCINDSRALDPDTLNTIKTHSLMDEVNPAFFSQPILYRT 292
QY 348 TS--S--LKGYHSSLPNRPRGKCLPDDQPIPTETFEYVADRNHPEVADQVEMGPILPFLH 403
Db 293 STIYRTQIAYDAQIKIRPGKTYVILVGTDHGKILIKSVAAESADSADKTVSVIEEDIV 352
QY 404 SK-YHYQKVAHV-RMQASHGETFVVLTLTDGRGIHK-V-VEEGE-QEH--S--FA-FNI 453
Db 353 LTKKEPIRNLEIYTMQIDPKGSDYDGKLIIVTSDQVAVALDLHRCHNDKITSCSECA 412

```

 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_p protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 3 16:24:21 1999; Maspar time 24.64 Seconds

Tabular output not generated. 574.766 Million cell updates/sec

Title: >US-09-240-410-2

Description: (1-666) from US09240410.pep

Perfect Score: 4968

Sequence: 1 MTRPPPGRAAPSAPRARVPG.....LAASIMLGVLPTLTGLLVH 666

Scoring table: PAM 150

Gap 11

Searched: 170751 seqs, 2126608 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

a-geneseq35
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 37.430; Variance 164.579; scale 0.227

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	661	13.3	771	13	R1380 Human semaphorin III	1.46e-48
2	620	12.5	477	13	R1415 Human collapsin.	1.01e-44
3	622	12.5	775	33	W63748 Human semaphorin.	6.59e-45
4	606	12.2	861	22	Mouse CD100 antigen.	2.09e-43
5	599	12.1	861	32	W58540 Human semaphorin.	9.30e-43
6	577	11.6	776	32	W51313 Rat semaphorin W.	1.05e-43
7	544	11.0	862	22	W17657 Human CD100 antigen.	1.23e-37
8	463	9.3	441	13	R1382 Vaccinia virus semaphorin	3.70e-30
9	459	9.2	730	13	R1379 Grasshopper semaphorin	8.62e-30
10	428	8.6	587	32	W51314 Human semaphorin W.	5.91e-27
11	390	7.9	974	33	W64221 Human secreted protei	1.68e-23
12	378	7.6	712	13	R1384 Tribolium semaphorin	2.04e-22
13	367	7.4	650	13	R1382 Drosophila semaphorin	2.00e-21
14	362	7.3	724	13	R1383 Drosophila semaphorin	5.63e-21
15	352	7.1	930	32	W57260 Human semaphorin Y.	4.45e-20
16	325	6.5	888	25	W19857 Human semaphorin Z.	1.14e-17

ALIGNMENTS

RESULT ID	1	Score	DB	Length	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
1	17	316	6.4	887	25	W19856	17	1.9e-17			
2	18	296	6.0	929	32	W57259	18	4.19e-15			
3	19	157	3.2	122	13	R1385	19	2.14e-03			
4	20	114	2.3	111	12	W51315	20	3.76e+00			
5	21	115	2.3	248	17	R87445	21	3.19e+00			
6	22	115	2.3	248	21	W09358	22	3.19e+00			
7	23	115	2.3	248	17	R96074	23	3.19e+00			
8	24	115	2.3	248	9	R46912	24	3.19e+00			
9	25	115	2.3	248	9	R55653	25	3.19e+00			
10	26	112	2.3	248	13	R67242	26	5.22e+00			
11	27	116	2.3	422	17	R87467	27	2.70e+00			
12	28	115	2.3	422	21	W09372	28	3.19e+00			
13	29	115	2.3	422	17	R96081	29	3.19e+00			
14	30	115	2.3	422	13	R67258	30	3.19e+00			
15	31	115	2.3	422	17	R86628	31	3.19e+00			
16	32	115	2.3	422	21	W09371	32	3.19e+00			
17	33	115	2.3	422	9	R46923	33	3.19e+00			
18	34	115	2.3	422	9	R55654	34	3.19e+00			
19	35	115	2.3	422	17	R87466	35	3.19e+00			
20	36	98	2.0	176	24	W14557	36	4.87e+01			
21	37	99	2.0	591	4	R23006	37	4.17e+01			
22	38	100	2.0	780	28	W47039	38	3.57e+01			
23	39	100	2.0	780	13	R62487	39	3.57e+01			
24	40	98	2.0	832	33	W61092	40	4.87e+01			
25	41	98	2.0	1091	27	W41641	41	4.17e+01			
26	42	98	2.0	1093	9	R56979	42	4.87e+01			
27	43	101	2.0	2329	25	W25038	43	3.05e+01			
28	44	96	1.9	832	33	W61090	44	6.64e+01			
29	45	95	1.9	2965	31	W56450	45	7.74e+01			

RESULT 1
 ID R1380 standard; Protein; 771 AA.
 AC R1380;
 DT 21-NOV-1995 (first entry).
 DE Human semaphorin III protein.
 KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
 KW variola major virus; smallpox; semaphorin receptor binding activity;
 KW modulation; nerve cell growth; immune response; viral pathogenesis;
 KW neurodegenerative disease; neuro-regeneration; oncological infection.
 OS Homo sapiens.
 PN W09507706-A.
 PD 23-MAR-1995.
 PE 13-SEP-1994; U10151.
 PR 13-SEP-1993; US-121713.
 PA (REGC) UNIV CALIFORNIA.
 PI Bentley DR, Goodman CS, Kolodkin AL, Mathes D;
 PI O'Connor T;
 DR WPI, 95-131177/17.
 DR N-PDB: Q87442.
 PT New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 PS Example 2: Page 60-63; 101pp: English.
 CC The sequence of the human semaphorin III protein.
 CC encoded by the grasshopper semaphorin III (087441), human semaphorin III,
 CC vaccinia virus semaphorin IV (087443), Drosophila semaphorin I and II,
 CC (087444-5), Tribolium semaphorin I (087446) or variola major (smallpox)
 CC virus semaphorin IV (087447) genes were used to generate a series of
 CC peptides (R70370-R70418), which retain semaphorin receptor binding
 CC activity. The semaphorin derived or semaphorin receptor derived peptides
 CC are potent modulators of nerve cell growth, immune responsiveness and
 CC viral pathogenesis. They can be used in diagnosis and treatment of
 CC neurological disease and neuro-regeneration, immune modulation and
 CC diagnosis and treatment of viral and oncological infection and diseases.
 SO Sequence 771 AA;
 Query Match 13.3%; Score 661; DB 13; Length 771;
 Best Local Similarity 30.3%; Pred. No. 1.46e-48;
 Matches 151; Conservative 126; Mismatches 174; Indels 47; Gaps 37;
 Db 166 rgksydpklltallsgelysgta-adfmgdfrlclghhprtqhsrwndp 224


```

QY      125 DCEVNTLLER-RSEGLIACGTNARHPSG--WNLVNGTVPLGEM-KGIYAFSPDSENSLV 180
D      167 vmvvgelysgts-ynfigsepisrnsbspilteyalpwlnepsfvadvigkspdpge 225
QY      181 LFEDEVYSTIRKOEYNGKIPFRFRIRGESSELYTSDV--MÖNQFIKATIVH-Q-D-Q 234
D      226 geddkvyfftevevevefeklmpvavayckdggglrtlg-kkwstflkarllcskp 284
QY      235 AYDKIYTFREDNPÖNPEAPLVNSVAQLCRDÖGGESSLVSXKMTFLKALVCSDA 294
D      285 dsglvfllldvflvrlap-gl-kepvfayftpqllnvgsavcaaylacveavfsrgky 342
QY      295 ATKNNFNRLÖDVFLLPDPSPGQMDTRVYGVFSNPMNW---SAYCVYSLGIDKVF-R-- 347
D      343 mgatveqgsttkwryngpvrtpgacidsaaraanytslnlpkcltqfvkdhplmd 402
QY      348 --TSSL---K-G-YHSSLPNRPCKCL-PDQÖPIP-TTFÖVADRNPE-V-AÖR-VE- 392
D      403 svrpidmrplkklkdvnyqivdrtgaldgtfydmfistdgalhkvallkevnh-v 460
QY      393 PMKPLKT-P-LFHSKHYHOKVAVHRMÖASHGETFHVLYLTDRGTHKVVPEGEÖHSFA 450
D      461 --ieetqlfrdeprvlltllsskkykfyasngsvvgaqlafcekhgs-ceedvland 517
QY      451 FNIMEIÖPFRRAAIIÖTMSIDAER-RKL-YVSÖWEVSÖVPLDCEVYGGCHGCLMSRD 508
D      518 PYCawapalkacvltlhgeas 538
QY      509 PYCGMDÖG-R-CISYSSERS 527

```

RESULT 6
ID W51313 standard; Protein: 776 AA.

AC W51313;
DT 08-SEP-1998 (first entry)
DE Rat semaphorin W.
KW Rat; semaphorin W; nerve extension inhibitor; antiallergic; anticancer;
KW Immunosuppressant; gene therapy; diagnosis; research reagent.
OS Rattus norvegicus.
PN M09815628-A1.
PD 16-APR-1998.
PF 03-OCT-1997; J03549.
PR 09-OCT-1996; JP-287636.
PA (SUMU) SUMITOMO PHARM CO LTD.
PI Kikuchi K, Kimura T;
DR WPI: 98-261015/23.
DR N-PSDB: V07279, V07280.
PT Nerve extension inhibitor protein semaphorin W - is useful as
PT therapeutic drug and diagnostic and research reagent
PS Claim 1; Page 60-64; 90pp; Japanese.
CC The present sequence represents rat semaphorin W. Semaphorin W and
CC its derivatives are nerve extension inhibitors which are useful as
CC antiallergic, immunosuppressant and anticancer agents. The DNA
CC encoding semaphorin W can also be used in gene therapy, e.g. using
CC a viral vector. The proteins, peptides, DNA and antibodies which
CC recognise the protein or peptides, can be used as diagnostic or
CC research reagents. Semaphorin W can be used as a screen for
CC semaphorin W antagonists with possible therapeutic use.
SQ Sequence 776 AA;

Query Match 11.6%; Score 577; DB 32; Length 776;
Best local similarity 27.8%; Pred. NO. 1,05e-40;
Matches 171; Conservative 149; Mismatches 240; Indels 56; Gaps 40;

```

D      5 aerprpgrppvpfipfppllllllaalsapcyrvpsrvptsjiseadsylrifa 64
QY      13 APRARV-PGPPARLGLPLRLRLLLMAAASÖGH-RSGPRI-FAVWKGHVQÖRDVDF 69
D      65 shlynyallvdpashtlylgardsifaltlpfsgerprridwmp-ethrgncrkkyk 123
QY      70 GÖTEPHTVLFHEPGSSVWVGGRGKYV-L-FDPEGKNMSV-RTVNIÖSTKÖCL--DKR 124

```

```

D      124 dechnfigllatvsnashlltctgtfcdpkydvsvsfqverlesgrgkcpfepaqrs 183
QY      125 -D-CENTITLLERSSE-LIACGTNARHPSÖMWL-YNG--TYVPLGEMGAYAFSPDSENS 178
D      184 aavmagvlylatvk-nfigteplisravgaedwrltelswlnapfvaaavspöe 242
QY      179 LVLFEGDEVYSTIRKOEYNGKIPFRFRIRGESSELYT-SDV--MÖNPÖFIKATIV-H-Q 232
D      243 wgedgddeffftetsrvlsyzerikvprvaycagliggrkrlg-grwtflkadll 301
QY      233 --DÖAVDDKITYFFFRDNDPKNPEAPLVNSVAQLCRDÖGGESSLVSXKMTFLKALV 290
D      302 cöpepögrsrylqmaelprpge-gtrpifgifsqvegaalsavcafrpqdlravln 360
QY      291 CSDAANKNFNRLÖDVFLLPDPSPGQMDTRVYGVFSNPMNW---YSAVCYVSLGIDKVF 347
D      361 gpfrelkhdcnrglpmwdehpqprpgecianmklqfgsgslspdrvtlfrdhplmd 420
QY      348 TS-SLÖK-GYHSSLP---N---PRGKCLPDQÖPIP-TTFÖVADRNPE-V-AÖR-VE 392
D      421 rpvfpadgrpllvtdctaylvvahrvtslsqkeydvlylgtedhlnravrigq-ls- 478
QY      393 -PMGPLKT-PLFHSK-YHYÖKVAVHRMÖASHGETFHVLYLTDRGTHKVVPEGEÖHSF 449
D      479 vledlal--fpepqvsesmklyhdw--llvshveitvynsnögrlqs-csecllaqdr 533
QY      450 AFNIMEIÖPFRRAAIIÖTMSIDAERKLYVSSÖWEVSÖVPLDCEVYGGCHGCLMSRD 509
D      534 vcawfrldaevahaghrngmqdlesadvsllcpkegphrvvfevpatvghvlpöcs 593
QY      510 YCGMD-Q-GRCISYSSERSVLOÖINPAEPHKECPNPRDPAKÖKVSÖLAPNSRYVLSCP 567
D      594 psawascvwhqpsgy 609
QY      568 MESRHATYSMRKENV 583

```

RESULT 7
ID W17657 standard; Protein: 862 AA.
AC W17657;
DT 24-JUL-1997 (first entry)
DE Human CD100 antigen.
KW CD100 antigen; semaphorin; leukocyte; B cell; T cell; lymphocyte;
KW vaccine.
OS Homo sapiens.
FH Key peptide
FT Location/Qualifiers
FT 1..41
FT /label= sig-peptide
FT 42..862
FT /label= Mat-protein
FT 42..553
FT /label= Semaphorin_domain
FT 554..630
FT /label= Ig-like_domain
FT 631..733
FT /label= stalk_domain
FT 735..752
FT /label= Transmembrane_domain
FT 753..862
FT domain
FT modified_site
FT 808..815
FT /label= Phosphorylation
FT /note= "putative tyrosine phosphorylation site"
PN M09717368-A1.
PD 15-MAY-1997.
PF 12-NOV-1996; U18645.
PR 09-NOV-1995; US-556422.
PA (DAND) DANA FARMER CANCER INST.
PI Boussoletic V, Freeman GJ, Hall KT, Nadler LM, Schultze JL;
DR WPI: 97-280962/25.
DR N-PSDB: T60665.
PT Nucleic acid molecule encoding CD100 antigen - which stimulates
PT leukocyte response, e.g. B cell aggregation, differentiation,

CC binding activity. The semaphorin derived or semaphorin receptor derived
CC peptides are potent modulators of nerve cell growth, immune
CC responsiveness and viral pathogenesis. They can be used in diagnosis and
CC treatment of neurological disease and neuro-regeneration, immune
CC modulation and diagnosis and treatment of viral and oncological infection
CC and diseases.
SQ Sequence 650 AA:

Query Match 7.4%: Score 367; DB 13; Length 650;
Best Local Similarity 28.4%: Pred. No. 2,00e-21;
Matches 124; Conservative 104; Mismatches 158; Indels 51; Gaps 38;

```
Db 3 dcdgnylrmvpspgrlfvcgtnsfmcmtylidsnlyleatkgavcypdphnst 62
Oy 125 DCENYITLLERSEGL-LACGTNARHPSCNN-LVNGTVPLGEMR-GYA--PSPDENS 179
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 63 svladmelysgtv-adfsgsdplylreplqtegydals-lnapnyf-ssft-qgd-f--- 114
Oy 180 VFEGDEVYSTIRKOEYNGKIPFRRIKGSELYTSDTWQNPQIKATIVHDDQAYDK 239
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 115 vyffretavefingckaly-srvarcvckwdyggphfr-nrwtcfkrlncsfpgdyp 172
Oy 240 IYFFREDNPD-KNEPAPLNVSRVAQLCRGDGSESSLSVKWNTFLKAMLVCSDAATNK 298
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 173 fyfneiqgsanlveggysmsklygvfnctpsnspgsavcafaiaqdadteggfkeq 232
Oy 299 -NFNRLQDVF-LLPDSGGMWRDTRYGVFSNPMN-Y--SAVCYSIGDI-DKV--FR-- 347
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 233 tglinsnwlpvnmakvpdpripgschndsrldpdtlnflkthlmdenvpafsfpllyrt 292
Oy 348 TS--S--LKGYSLSLPNPRPGKCLPDQOPIPETFOVADRHEVAQRVEPMPLTPLEH 403
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 293 stlyfctglavdaqlktpgsklydvlfygtbqklyknsaasadsadvtvsvleaidv 352
Oy 404 SK-YHYQKAAVH-RMQASIGETFWLYLTDDKGTIHK-V-VEPGE-QEH--S--FA-FNI 453
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 333 ltksepitnleivrtmtygqpkdgsyddgkllivdsqvaiaqprhndkittsecvca 412
Oy 454 M-EIQPFRRAAAIQTMSLD-A-E-R--R-KLYVSSQWEVSQVPLDLCEYVG-GGCHGLM 505
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 413 lqdyvcawdklagkcrs 429
Oy 506 SRDPYCGWDO--GRCIS 520
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 14
ID R71383 standard; Protein; 724 AA.
AC R71383;

DE 21-NOV-1995 (first entry)
DT Drosophila semaphorin II protein.
KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
KW varicella major virus; smallpox; semaphorin receptor binding activity;
KW modulation; nerve cell growth; immune response; viral pathogenesis;
KW neurological disease; neuro-regeneration; oncological infection.
OS Drosophila sp.
PN WO9507706-A.
PD 23-MAR-1995.
PF 13-SEP-1994; U10151.
PR 13-SEP-1993; US-121713.
PA (REGC) UNIV CALIFORNIA.
PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
PI O'Connor T;
DR N-PSDB: 087445.
DR N-PSDB: 087445.
PT New class of semaphorin peptide(s) and polypeptide(s) - are
PT potent modulators of nerve cell growth and regeneration
PS Example 2; Page 79-82; 10pp; English.
CC The sequence of the Drosophila semaphorin II protein. The gene was
CC isolated by PCR using primers based on sequence homology between the
CC complete grasshopper and partial Tribolium and Manduca (moth) semaphorin
CC gene sequences. The products gave two different sequences, each of which
CC was used to obtain its respective complete sequence: semaphorin I
CC (087444) and II. The proteins encoded by the grasshopper semaphorin I

CC (087441), human semaphorin III (087442), vaccinia virus semaphorin IV,
CC Drosophila semaphorin I and II, Tribolium semaphorin I (087446) or
CC varicella major (smallpox) virus semaphorin IV (087447) genes were used to
CC generate a series of peptides (R70370-R70418), which retain semaphorin
CC receptor binding activity. The semaphorin derived or semaphorin receptor
CC derived peptides are potent modulators of nerve cell growth, immune
CC responsiveness and viral pathogenesis. They can be used in diagnosis and
CC treatment of neurological disease and neuro-regeneration, immune
CC modulation and diagnosis and treatment of viral and oncological infection
CC and diseases.
SQ Sequence 724 AA:

Query Match 7.3%: Score 362; DB 13; Length 724;
Best Local Similarity 29.0%: Pred. No. 5.63e-21;
Matches 84; Conservative 68; Mismatches 113; Indels 25; Gaps 19;

```
Db 265 vyffretavefingckavy-srvarcvckdvygknlla-hmatalylarlnsisgefp 322
Oy 240 IYFFREDNPD-KNEPAPLNVSRVAQLCRGDGSESSLSVKWNTFLKAMLVCSDAATNK 298
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 323 fyfneiqsvyqpsdksrffat-ftstnglgsavcsfhlneigaafngkfkegssn 380
Oy 299 -NFNRLQDVFLLPDSGGMWRDTRYGVFSNPMNLSAVCVSYSGDID--KV-FR--RS- 349
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 381 sawlpvlnsrpeprgtcvndtsnldvtvlnflrshplmdkavnh-eh-npyykrdl 438
Oy 350 S--LKGYSLSLPNPRPGKCLPDQOPIPETFOVADRHEVAQRVEPMPLTPLEH-X 406
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 439 vftkiavckitldlmglylyyyvgtnlgrlyklyqyyngesls-klldifevapneal 497
Oy 407 HYQKAAVHMQAS-HGEFFHWLYLTDDKGTIHKVPEDEQHSFAFNIETOPFRRAAI 465
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 498 qvmeisqrklslygtidhrklyqldiamcnrrydh-cfrc-v-rdpbgcwg 544
Oy 466 QTMSLDARRRLIYSSQWEVSQVPLDLCEV-YGGGCHGLMSRDPYCGMD 514
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 15
ID W57260 standard; Protein; 930 AA.
AC W57260;

DT 02-SEP-1998 (first entry)
DE Human semaphorin Y.
KW Human; semaphorin Y; nerve extension inhibitor; central nervous system;
KW peripheral nerve growth.
OS Homo sapiens.
PN WO9811216-A1.
PD 19-MAR-1998.
PF 09-SEP-1997; J03167.
PR 08-AUG-1997; JP-227220.
PR 11-SEP-1996; JP-263565.
PA (SUMU) SUMITOMO PHARM CO LTD.
PI Kikuchi K, Kimura T;
PI WPI: 98-250958/22.
DR N-PSDB: V28915, V28916.
PT DNA encoding human and rat semaphorin Y - an inhibitor of nerve
PT extension
PS Claim 1; Page 65-70; 85pp; Japanese.
CC The present sequence represents human semaphorin Y which inhibits
CC nerve extension. Semaphorin Y genes and proteins may be used to inhibit
CC peripheral nerve growth. Semaphorin Y antagonists can be used to
CC accelerate regrowth of the central nervous system.
SQ Sequence 930 AA:

Query Match 7.1%: Score 352; DB 32; Length 930;
Best Local Similarity 28.4%: Pred. No. 4.45e-20;
Matches 132; Conservative 112; Mismatches 164; Indels 56; Gaps 44;

```
Db 120 ecnyatrvlpwsgtllacgtnsfpcrsgyitlqgegeelsqgarcpidatsnva 179
Oy 125 DCENYITLLERR-SEGLACGTNARHPSCNNLVNGTVPLGE-MRGYA--PSPDENS 180
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 180 lfeegslyata-adfqsadavvyslqpppl-ts-a-kygskwlrephfygalehghn 235
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```


This Page Blank (uspto)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Masrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 3 16:30:18 1999; Maspar time 39.74 Seconds

Tabular output not generated. 914.651 Million cell updates/sec

Title: >US-09-240-410-2

Description: (1-666) from US09240410.pep

Sequence: 1 MTPPPGRAPSAAPRAVP...LAASIMGLVLTITGLLVH 666

Scoring table: PAM 150

Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle
9:sp.phage 10:sp.plant 11:sp rodent 12:sp.unclassified
13:sp.vertibrate 14:sp.virus

Statistics: Mean 49.772; Variance 78.204; scale 0.636

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	4968	100.0	666	4	SEMAPHORIN L.	0.00e+00
2	2595	52.2	393	11	SEMAPHORIN L. (FRAGMENT	0.00e+00
3	1907	38.4	653	14	SIMILAR TO GENBANK ACC	0.00e+00
4	737	14.8	748	11	SEMAPHORIN A PRECURSOR	1.22e+150
5	725	14.6	751	11	SEMAPHORIN A PRECURSOR	1.53e+147
6	716	14.4	749	4	SEMAPHORIN V.	3.21e+145
7	716	14.4	750	4	SEMAPHORIN E.	3.21e+145
8	710	14.3	751	4	SEMAPHORIN E.	1.13e+143
9	705	14.2	751	13	COLLAPSED 3	2.19e+142
10	688	13.8	753	4	SEMAPHORIN III FAMILY	5.20e+138
11	685	13.8	772	13	COLLAPSED.	3.07e+137
12	682	13.7	754	11	SEMAPHORIN IV ISOFORM	1.81e+136
13	682	13.7	785	11	SEMAPHORIN IV ISOFORM	1.81e+136
14	681	13.7	785	4	SEMAPHORIN IV.	3.27e+136
15	679	13.7	785	4	SEMAPHORIN.	1.07e+135
16	669	13.5	772	11	SEMAPHORIN D PRECURSOR	3.94e+133
17	668	13.4	772	11	SEMAPHORIN III/COLLAPS	7.11e+133
18	668	13.3	771	11	SEMAPHORIN III.	4.43e+131
19	628	12.6	775	4	SEMAPHORIN H.	1.22e+122
20	618	12.4	782	11	SEMAPHORIN C (SEM C) (4.35e+120

21	599	12.1	861	11	SEMAPHORIN J (SEMAPHOR	2.99e+115
22	594	12.0	761	13	COLLAPSED-2.	5.57e+114
23	597	12.0	775	4	KIAA0331.	9.62e+115
24	572	11.5	785	13	COLLAPSED 5.	2.11e+108
25	568	11.4	294	13	COLLAPSED-3 (FRAGMENT)	2.18e+107
26	544	11.0	862	4	SEMAPHORIN.	2.50e+101
27	522	10.5	834	11	SEMAPHORIN I (M-SEMA F	8.52e+96
28	511	10.3	299	13	COLLAPSED-5 (FRAGMENT)	4.88e+95
29	480	9.7	760	11	SEMAPHORIN B PRECURSOR	2.67e+85
30	473	9.5	295	13	COLLAPSED-4 (FRAGMENT)	1.46e+83
31	461	9.3	730	5	FASCICLIN IV.	1.38e+80
32	448	9.0	1074	4	SEMAPHORIN F HOMOLOG.	2.26e+77
33	441	8.9	1077	11	SEMAPHORIN F PRECURSOR	1.20e+75
34	416	8.4	1093	11	SEMAPHORIN G PRECURSOR	1.65e+69
35	390	7.9	494	4	SEMAPHORIN F (FRAGMENT	3.52e+63
36	392	7.9	888	11	SEMAPHORIN VIA.	1.15e+63
37	376	7.6	712	5	SEMAPHORIN-I PRECURSOR	2.81e+60
38	367	7.4	562	5	CESEMA.	1.25e+57
39	367	7.4	771	5	SEMAPHORIN-I.	1.25e+57
40	362	7.3	706	5	SEMAPHORIN-II.	1.99e+56
41	341	6.9	284	11	SEMAPHORIN IV HOMOLOG	2.05e+51
42	316	6.4	886	11	SEMAPHORIN N (SEMAPHOR	1.63e+45
43	316	6.4	887	11	SEMAPHORIN Z.	1.63e+45
44	283	5.7	770	5	SIMILAR TO SEMAPHORIN-	7.46e+38
45	204	4.1	1963	4	KIAA0463 PROTEIN (FRAG	2.42e+20

ALIGNMENTS

RESULT ID	1	PRELIMINARY;	PRT;	666 AA.
AC 075326;				
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)				
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)				
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)				
DE SEMAPHORIN L.				
GN SEMAL.				
OS HOMO SAPIENS (HUMAN).				
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;				
OC CATARRHINI; HOMINIDAE; HOMO.				
RN [1]				
RP SEQUENCE FROM N.A.				
RP MEDLINE: 98389619.				
RA LANGE C., LIEHR T., GOEN M., GEEHART E., FLECKENSTEIN B., ENSSER A.;				
RT "New eukaryotic semaphorins with close homology to semaphorins of DNA				
RL GENOMICS 51:340-350(1998).				
DR EMBL: AF030698; G3523115; -				
SO SEQUENCE 666 AA; 74823 MW; F546B39D CRC32;				

Query Match	100.0%	Score 4966;	DB 4;	Length 666;
Best Local Similarity 100.0%;	Pred. No. 0.00e+00;	Mismatches 0;	Indels 0;	Gaps 0;
Matches 666; Conservative				
Db 1	MTPPPGRAPSAAPRAVP	GPAPRLGLPLRLILLMAAASNOGHLRSGPRIFAYWKG	60	
Qy 1	MTPPPGRAPSAAPRAVP	GPAPRLGLPLRLILLMAAASNOGHLRSGPRIFAYWKG	60	
Db 61	HVGODRVDFGQTEPHYVLF	HEPSSSVWVGGRKVVLFDEPGRKMSVRTVNGTSGSC	120	
Qy 61	HVGODRVDFGQTEPHYVLF	HEPSSSVWVGGRKVVLFDEPGRKMSVRTVNGTSGSC	120	
Db 121	LDRDCENTYTLLEBRSEGL	LACGTNARHPSCHNLVNGIVYVPGEMRGVAFSPDENSIV	180	
Qy 121	LDRDCENTYTLLEBRSEGL	LACGTNARHPSCHNLVNGIVYVPGEMRGVAFSPDENSIV	180	
Db 181	LFEGDEVYSTIRKOEYNGKIP	FRFRIRGESELYTSDVVMONPOFIKATVHOQAYDCKI	240	
Qy 181	LFEGDEVYSTIRKOEYNGKIP	FRFRIRGESELYTSDVVMONPOFIKATVHOQAYDCKI	240	
Db 241	YFFREDNPDKNDEAPLNV	SRVAOLCRDGGGESSLSVSKWNTFLKAMLVCSDAATKNF	300	
Qy 241	YFFREDNPDKNDEAPLNV	SRVAOLCRDGGGESSLSVSKWNTFLKAMLVCSDAATKNF	300	

Db	301	NRLDVLFLDPBSQWMDTVYGVENPNMNYSAVCYYSLGJLIDKXFRSSLKGYHSSLPN	360
Qy	301	NRLDVLFLDPBSQWMDTVYGVENPNMNYSAVCYYSLGJLIDKXFRSSLKGYHSSLPN	360
Db	361	PRPGKCLPDDOPIPTETFFQVADRHPRVPAORVEMPGPLKPLFHSYHYHOKAVYRMASH	420
Qy	361	PRPGKCLPDDOPIPTETFFQVADRHPRVPAORVEMPGPLKPLFHSYHYHOKAVYRMASH	420
Db	421	GEIFHVLYLTDRGTLHKVVEPGEQESHFAFIMEIOFRRAAAIQTMSLDAERKLYVS	480
Qy	421	GEIFHVLYLTDRGTLHKVVEPGEQESHFAFIMEIOFRRAAAIQTMSLDAERKLYVS	480
Db	481	SQWESVQVPLDLEVGYGCGGCHGLMSRDYCGMDGRCISYSSRSVLOSINPAEPHKE	540
Qy	481	SQWESVQVPLDLEVGYGCGGCHGLMSRDYCGMDGRCISYSSRSVLOSINPAEPHKE	540
Db	541	CPNPKPKAPLOKVS LAPNSRYTLC PMESNRATY SWRHKENVEDSCBPGHOSNCLFTI	600
Qy	541	CPNPKPKAPLOKVS LAPNSRYTLC PMESNRATY SWRHKENVEDSCBPGHOSNCLFTI	600
Db	601	ENLTAQOYGYHYFCEAEGSYFREDQHMQLLPEDGIMAEHLGLHACALAASTMLGVLPRTL	660
Qy	601	ENLTAQOYGYHYFCEAEGSYFREDQHMQLLPEDGIMAEHLGLHACALAASTMLGVLPRTL	660
Db	661	LGLLVH 666	
Qy	661	LGLLVH 666	

RESULT	2			
ID	O88371	PRELIMINARY:	PRI:	393 AA.
AC	O88371;			
D8	01-NOV-1998 (TREMBLREL. 08,			CREATED)
DT	01-NOV-1998 (TREMBLREL. 08,			LAST SEQUENCE UPDATE)
DT	01-NOV-1998 (TREMBLREL. 08,			LAST ANNOTATION UPDATE)
DE	SEMAPHORIN L (FRAGMENT).			
GN	SEMAL.			
OS	MUS MUSCULUS (MOUSE).			
OC	EDUAROTIA; METAEOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;			
OC	SCUROGNATHI; MORIDAE; MORINAE; MUS.			
OC	[1]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE: 98389619.			
RX	LANGUAGE C.; LIEHR T.; GOEN M., GEBHART E., FLECKENSTEIN B., ENSER A.;			
RA	"New eukaryotic semaphorins with close homology to semaphorins of DNA			
RT	viruses.";			
FT	GENOMICS 51:340-350(1998).			
DR	EMBL: AF030699; G3523117; -.			
SO	NON TER 393			
FT	SEQUENCE 393 AA; 43909 MW; 7CB8193C CRC32;			
	Query Match	52.2%; Score 2595; DB 11; Length 393;		
	Best Local Similarity	90.2%; Pred. No. 0.00e+00;		
	Matches 357; Conservative	18; Mismatches 17; Indels 4; Gaps 2;		
D8	1 MTPEPPGRAPASARAAVLSPAFEGPLRLRLILVWVAASAQGSRSRGPRISAVWKG 60			
QY	1 MTPEPPGRAPASARARVPGPAPRLGLRLRLLLMAAASQGLRSGPRIFAVWKG 60			
D8	61 --ODHWDFSQPEPHIVLFHPGFSFVWGVRGVYHFNFPEEKNASVTRVINIGSTGSC 117			
QY	61 HVGDGDVDFGGTERTHTLVLFHPGSSSVWVGGRGVLYLFDPPECKNASVTRVINIGSTGSC 120			
D8	118 QDKDDCGNYITTLERNGNLVCGTNARKPSCMWLVNDVSVMSLGEMKGYAPFSPDENSL 177			
QY	121 LDKRDCENYITTLERRESEGLAGCTNARNHSCMWLVNGVTYV-ELGEMRGYAPFSPDENSL 179			
D8	178 VLEFGDEVSTIRKOENGKIIPRRIRLRGESELYTSPTWMONOQFIKATIVHDOADVDK 237			
QY	180 VLEFGDEVYSTIRKOENGKIIPRRIRLRGESELYTSPTWMONOQFIKATIVHDOADVDK 239			
D8	238 IYVFREDNDKNEADPLNVSRVAQLCRGDGGESSLVSKNWTFLKAMLYCSDAATNRN 297			

QY	240	IYFFREDNPDKNEAPLANSRYAOLCRRDDGGESSLSYKKNFTFLKMLYCSDAALNTNK	299
Db	298	FNRLQDYFLLPDSGGQREDTRYGVFSNPMWNTSAVCVYSLGQIDBVFETSSLKQYHMGLS	357
QY	300	FNRLQDYFLLPDSGGQREDTRYGVFSNPMWNTSAVCVYSLGQIDBVFETSSLKQYHMGSLP	359
Db	358	NPRGMLPKPKKQPIPTETFGVADSHFEVAKORPEPG	393
QY	360	NPRGKCLPDDQPIPTETFGVADSHFEVAKORPEPG	395

```

RESULT      3
ID          064906      PRELIMINARY;      PRT;      653 AA.
AC          064906;
DT          01-NOV-1996 (TREMBLREL. 01, CREATED)
DT          01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT          01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE          SIMILAR TO GENBANK ACCESSION NUMBER L26081.
OS          ALCELAAPHINE HERPESVIRUS 1.
OC          VIRUSES: DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE:
OC          GAMMAHERPESVIRINAE.
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN-CS00;
RC          MEDLINE; 97201573.
RA          ENSNER A., FLECKENSTEIN B.;
RT          "Alcelaphine herpesvirus type 1 has a semaphorin-like gene.";
RL          J. GEN. VIROL. 76:1063-1067(1995).
RN          [2]
RP          SEQUENCE FROM N.A.
RC          STRAIN-CS00;
RC          MEDLINE; 97404659.
RA          ENSNER A., FFLANZ R., FLECKENSTEIN B.;
RT          "Primary structure of the alcelaphine herpesvirus 1 genome.";
RL          J. VIROL. 71:6517-6525(1997).
RN          [3]
RP          SEQUENCE FROM N.A.
RC          STRAIN-CS00;
RC          ENSNER A., FFLANZ R., FLECKENSTEIN B.;
RL          SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR          EMBL; U18243; GI000717; -.
DR          EMBL; AF005370; G2337970; -.
SQ          SEQUENCE      653 AA;  73645 MW;  CCB53C9 CRC32;

```

[illegible]

Db	401	QVARRPEVADPVQYKXNNAMFPIQSKYTKKLVYRREY-GVFMAFIPLTKIGTH	459
Db	401		
Qy	379	QVARRPEVADPVQYKXNNAMFPIQSKYTKKLVYRREY-GVFMAFIPLTKIGTH	437
Db	460	IYVRVEDNSTALVILTEINFQKPAFIONLINTNLKLVNSEWSEVPLDLCYVG	519
Qy	438	KVREGDEHSEFAFIMETIOFFRRRAAQTMSLDAERKLVSSQWEVSQVPLDLCYVG	497
Db	520	NDCESEMSRDPPLCTWYNNCT-S-FK-QRVSEVETGPPANRLTSEMCGDHAPVYKHOVS	576
Qy	498	GGCHGLMSRDPYCGMDGRCISITSSRSVLTQSLNPAEPH-KE-CPRKRDKAPLQXVS	555
Db	577	IPLISNYSLSCPAYSNAHDYFWYTDGFEKCKHYKTKHNCDCILIANSTFNTGHCNM	636
Qy	556	LAPNSRYLSCPMSERHATYSWRHKENEQSCCEPGHOSPNCLFLENTLAQOYGHYCEA	615
Db	637	KEDS 640	
Qy	616	QEGS 619	
RESULT 4			
ID	062177	PRELIMINARY;	PRT; 748 AA.
AC	062177;		
D8	01-NOV-1998 (TREMBLREL. 08, CREATED)		
D7	01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)		
D7	01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)		
DE	SEMAPHORIN A PRECURSOR (SEM A).		
GN	SEMAA OR SEMA.		
OS	MUS MUSCULUS (MOUSE).		
OC	EDUAROTIA; METAEOA; CHORADIA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;		
OC	SCUROGNATHI; MORIDAE; MORINAE; MUS.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NMRI:		
RX	MEDLINE; 95267431.		
RA	PUDSHEL A.W., ADAMS R.H., BETZ H.;		
RT	"Marine semaphorin D/collapsin is a member of a diverse gene family		
RL	and creates domains inhibitory for axonal extension.";		
CC	NEURON 14:941-948(1995).		
CC	-1- FUNCTION. INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO		
CC	SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.		
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR (PROBABLE).		
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO.		
CC	LOW LEVELS FOUND BETWEEN DAYS 10-12.		
CC	EXPRESSION PEAKS ON DAY 13 WITH MODERATE LEVELS FROM THEN UNTIL		
CC	BIRTH.		
CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.		
CC	CONTAINS ONE C2-LIKE DOMAIN.		
DR	EMBL: X85990; G854324; -.		
DR	MED: MG1:107561; SEMAA.		
DR	PFAM: PF00047; 19; 1.		
KW	SIGNAL. IMMUNOGLOBULIN FOLD; MULTIGENE FAMILY; NEUROGENESIS;		
FW	DEVELOPMENTAL PROTEIN.		
FT	SIGNAL 1 26	POTENTIAL.	
FT	CHAIN 27 748	SEMAPHORIN A.	
FT	DOMAIN 586 649	IG-LIKE C2-TYPE DOMAIN.	
SO	SEQUENCE 748 AA; 82894 MW; A7E53A8D CRC32;		
Query Match 14.8%; Score 737; DB 11; Length 748;			
Best Local Similarity 33.7%; Pred. No. 1,22e-150;			
Matches 168; Conservative 112; Mismatches 166; Indels 52; Gaps 36;			
Db	112	ECMNPVRLAHYNTHTLACRTGAFFPTCALMRNATAGGTHASGPERKLEDGKGTYPD	171
Qy	125	DCENTILTLERSRG-LIACGTNARHPC---NNLVNGT-V-V-P-LGEMRGYAPSP	174
Db	172	RHRPVSVLGEELYSGV-TADLMGRDFTIFRSLGONPSLRTPEHDSRWLNPKRVKFWI	230
Qy	175	DENSLVLFEGDEGVSTIKQYENCK-IRPRFRIGESLRY-S-DIV-MQNPQIKAT-I	229
Db	231	PESENDDDKITFFEFRESAVEAEPAMGMSYSRVQICRNDLQGOQRL-VKKWTFELKAR	289

[illegible]

Oy	227	ATIHQD-QAVADKITYFFEFNDPNKAPNPALNVSVAVQJLDCGDDGGSSLSVSKMNTFL	265
Db	285	KARLYCVSTDEDDPEHFEDELEVFL-ETDNP-RTLYLGGFTTSSSVFKGSAVCYHL	342
Oy	286	KAMLYCS-PA-ATNKNFNLQDVFLLPBPSCQWMDRTRYGVFSNPNM-Y-SAVCYSL	339
Db	343	SDIOITVNGFPANKEGNHQLISYOGIRIYPGPGTCGGAFPPNNRTTKDPPDVVFIR	402
Oy	340	GDIDKVRITS-SLK-G-----YHSLNPNRPKGLCPQ-QP-I-PIETI-QVADRRP	365
Db	403	NHPLMVISPIHRHRLVIRIGTDYKYTKIADVRYNAADG-RHYHLFGTDGTVQKVV	461
Oy	386	EVAARVPMPRL-KTLPF-H-SKYHYQVAAVHRMASHGETFHVLYLTTRGTIRHVE	441
Db	462	LPTSSASGELLLEELVEFNHNPILTMTISSKKOOLVYSSNEGVSQVSLRHCIYTAG	521
Oy	442	-PGEQHSFAPNTIMEIOPEFRRAAIOTMTSLDERKRLTVSSQWESQVOPDLCEVYGGC	500
Db	522	ADCCLAADPCYAMDGHSCSRFPYTGRRSRRODVRHGNPLQCRFNKLKARNAAEIYQY	581
Oy	501	HGCLMSRDPCGMDQGRGCIISYSE-RSVLDSINPADRHNECP-NKPRKALQKVS	556
Db	582	GVRNNSTFLECAKPSPOASIKTLLQDK	609
Oy	557	A-PNSRYLSCPMESRRATYSMR-NHEN	582

RESULT	6		
ID	Q13214	PRELIMINARY:	PRT: 749 AA.
AC	Q13214;		
DT	01-NOV-1996 (TREMBLREL, 01, CREATED)		
DT	01-NOV-1996 (TREMBLREL, 01, LAST SEQUENCE UPDATE)		
DT	01-NOV-1998 (TREMBLREL, 08, LAST ANNOTATION UPDATE)		
DE	SEMAPHORIN V,		
OS	HOMO SAPIENS (HUMAN).		
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES		
OC	CATARRHINI; HOMINIDAE; HOMO.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE: 96210603.		
RA	SKETIO Y., BADER S., LATIF F., CHEN J.Y., DUH F.M., WEI M.H.,		
RA	ALBANDSI J.P., LEE C.C., LEMMAN M.I., MINNA J.D.;		
RT	"Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung		
RT	cancer deletion region and demonstrate distinct expression		
RT	patterns.";		
RL	PROC. NAT'L. ACAD. SCI. U.S.A. 93:4120-4125(1996).		
DR	EMBL: U28369; G974284; -.		
DR	PERM: F00047; I9; 1.		
QO	SEQUENCE 749 AA; 83121 MW; 6138C5D5 CRC32;		

	14.4%;	Score 71.6;	DB 4;	Length 749;
Query Match	Similarity 33.0%;	Pred. No. 3,216/145;		
Best Local Match	167;	Conservative 115;	Mismatches 170;	Indels 54;
				Gaps 41.
Db	112	ECMNFVKLLAHAYNRTLHLCAGTGAFHPYCAVEVGHRAEPEVRLRLDPGRILEDGKGSYPD	171	
	:	:	:	:
	:	:	:	:
	:	:	:	:
Qy	125	DCENTITLLER-RSBSGLIACGTNRHPSG-WNLV-N-G--IVLVL--G--EM-RGIAPFS	173	
Db	172	PRHRAASVLYGEEVLISGYA--ADLMGRDFTIFRSIGQRPSELTPEHDSRWLNPEKFAVFW	230	
	:	:	:	:
	:	:	:	:
	:	:	:	:
Qy	174	PDENSLVLFEEDEDEVSTIRKOEYNGK-IPRRRRIRGESELYT--DMV-MQNPFIKAT-	228	
Db	231	IPESNPDDDKIYFEFRETVAEAPALGRLSVSVGQICRNDVGGQSL--VNKKTITLKA	289	
	:	:	:	:
	:	:	:	:
	:	:	:	:
Qy	229	IVHQQAIVDDKIYFERDNPDKNPE-APLWVSRAQICRQDGGSESLYSKWNTFLKA	287	
Db	290	RLVCSVPGEVDTHFDQLQDVFLLS-SRDH-RTPLLYAVFSTSSIFQGSAAVCYSMDV	347	
	:	:	:	:
	:	:	:	:
	:	:	:	:
Qy	288	MLVCS--DAATINKNNRLLQDVFLLPDPGQQRDRIRYVGSFNPNV-N-Y--SAYCYISLGI	342	
Db	348	RRAFLGPAHHEGPHHWAISYQGRVLYPRQMCSSKITGTSTSSIKDDPDVDYIOFARNHPL	407	
	:	:	:	:
	:	:	:	:
	:	:	:	:

QY 343 DK-V-P-RTSSL-K-GVHSLPMPRRGCKLPD-QP-I-P-Terf--QVAD-R-HEE 386

Db 408 MINSVLPPLGG-R-PLFTLOVCANITFTTOIAADRYAABDGH-YDVLFTGTDVGLVKAISVP 464

QY 387 VAOREPMPGPKLPTEL-F-H-SKYITQVAAVHMQASHGCFTHLYLTTLTGTHIKVVE-P 442

Db 465 KGSRPSAEGLLLEELHVEEDSAVTSQMOSIKSRHOLYVASRSVAOIALHRCANHRVCT 524

QY 443 -GEQDHSFAFIMTLOFPFRRAAAIQTMSLAERKRLIYSSQMEVSOVPLDLCVYTGCGH 501

Db 502 ECLCLADPYCAMDCACTRFQPSAKRRFRBODVYRNGDPSTLCSGDSRPALLLEHKYGV 584

QY 502 GCLMSRDYCCMGDGRCSIIYSRSYVL-QSINAPPEHKECENPRPKAPL-QKVSAP 558

Db 585 GSAFLCECPRLQARVEMTFQKRGV 610

QY 559 NSRYILSCPMESRHATYSWR-HKENV 583

```

RESULT      7
ID          Q93018      PRELIMINARY;      PRT;      750 AA.
AC          Q93018;
DT          01-FEB-1997 (TREMBLREL. 02, CREATED)
DT          01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT          01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE          SEMAPHORIN V.
OS          HOMO SAPIENS (HUMAN).
OC          EUDAROTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; PRIMATES;
OC          CATARRHINI; HOMINIDAE; HOMO.
[1]
RN          SEQUENCE FROM N.A.
RA          DANTE M., WAMSLEY P.;
RL          SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RN          SEQUENCE FROM N.A.
RA          WATERSTON R.;
RL          SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
[3]
RN          SEQUENCE FROM N.A.
RA          WATERSTON R.;
RL          SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL: U73167; G2880035; -.
PFAM: PF00047; 19; 1.
QO          SEQUENCE      750 AA;      83034 MW;      C8C87E34 CRC32;

```

Query Match	14.4%	Score 716	DB 4	Length 750
Best Local Similarity	33.0%	Pred. No. 3.21e-145		
Matches 167	Conservative 115	Mismatches 170	Indels 54	Gaps 41
Db	113	ECMNFVKLLAHYNTTHLACGTGAHFPTCAFEVGHNAEEBPVLRDLDPGRITEDGKGSYPD	172	
Qy	125	DCENITITLLER-RSEGLLCATNRRHSC-WNLV-N-G--TVVPL--G--EM-RGIAPFS	173	
Db	173	PRHRAASYLVEELLYSGVA-ADLMGRDFTIFRSLGQRPSSLRTPEHDSRWMLNEPKVFW	231	
Qy	174	PDENSLVFEEDDEVYSTIRKQYNGK-IPRRRIIRGESELYT-S-DTV-MQNQFIKAT-	228	
Db	232	IPSENPDDDKIYFFFRFETANVEAPALGRSLSVNGOICRNDVGQNSL-VNKKTTFLKA	290	
Qy	229	IVHDDAIDDKIYFFFRDNPDKNPE-APLWSSVPAOLCGDGGESSLSVSKWTFLLKA	287	
Db	291	RLVCSVPEVEDDTEFDLOLQVFIILS-GRDH-RPTEILAVRSTSSIRQGSVCVYSKNDV	348	
Qy	288	MLVCS--DAANKNFENRLQDVFIILPDSGCKRQIRVYGVSNPNV-Y--SACVYSISGDI	342	
Db	349	RRALFLGPAHKEGEMHOMVSYQGVVPPRPMCSKTFGTSSITKDPDDVIOFARNHPL	408	
Qy	343	DK--V--F--RTSSL-K--GVHSSLPMRPCKLPDQ-QPIL-TEFE--QVAD--R-RPE	386	
Db	409	MYNVSLPRIGS-R-PLFLQVGNATYFTLOIADRAVAADGH-YDVLFITDGVTLKATVSP	465	
Qy	387	VAOREEPGPKLTFL-H-SKYVQYQVAVHARMDASGEMFHLVLLTIDTGTLHKVE-P	442	

```
Db 466 KGRSAGELLEEHLVEDESAVTSMOISSKRLQLYVASRGAVALHRCACAGRCT 525
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 443 -GEQHSFAFNIMEIQPRRAAIIOTMSIDARRLIYSSQWESQVPLDCEYVGGCH 501
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 526 ECLLARDYCAMDGYACTRFOPSAKRFRRODVNRGDPSTLCSGSSRPALLEHVKFGE 585
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 502 GCLMSRDPYCGWDGRCISITYSSESVL--QSLNPAEPHKECPNPKPAFL-QKVS LAP 558
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 586 GSSALFCEPRSLQARVEKTEFORAGY 611
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 559 NSRYLSCPMESRAHATYSWR-HKENV 583
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 8 PRELIMINARY: PRT: 751 AA.
ID 099985 AC 099985:
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SEMAPHORIN E.
OC HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA YAMADA T., ENDO R., GOTOH M., HIROHASHI S.;
RT "Identification of semaphorin E as a non-MDR drug resistance gene of
human cancers."
RL PROC. NATL. ACAD. SCI. U.S.A. 94:14716-14718(1997).
DR EMBL: AB000220; D1033360; -.
SQ SEQUENCE 751 AA; 85207 MW; F89FB934 CRC32;

Query Match 14.3%; Score 710; DB 4; Length 751;
Best Local Similarity 33.6%; Pred. No. 1,13e-143;
Matches 151; Conservative 104; Mismatches 151; Indels 43; Gaps 30;

Db 167 SENPNVNTSVYVINEELFSGMYI-DFMGTDALFSLTKRNAVVRTDQHSKWLSEPIFVD 225
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 171 PFSPEPNSLVLFEGDEYVSTIRKQYNGK-IPFRFRIGSESELYTS--DTV-MQNPQFIK 226
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 226 AHVDPGTDPNDAKYFFFEKRLIDNNRSTKQISHMARIKCPNDTGGLSL-VNKNWTFEL 284
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 227 ATIVHOD-QAYDDKIYFFREDNPKNEAPLNSRVAQLCRGDGGESSLVSXWNTFL 285
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 285 KARLYCSVTDEDEGPETHEDVELL-ETDNP-RTILVYGIFTTSSSVFKGSACVYHL 342
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 286 KAMLYCS--DA-ATKKNFNRLQDVFLPDPSCQWDRTRYGVFSNPMN-Y--SACVYSL 339
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 343 SDIQVFNGPFAHKEGPNHQLSTYQGRIPYRPGTCPGGAFTPNMRTTEFPDDVYTFIR 402
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 340 GDIDVFRTS-SLK-G-----YHSSLPNRPKGLPQ-QP-IPTEF-F--QVAD-R 383
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 403 NHPLMYNSIYPIHKRPPLIVRI-GTDYKYTKIADVRYNADG-RYHFLFGTDRGVQKVV 460
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 384 -HPEVAQREPM-G-PLKPRFLHSHYKQVAVHMQASHGETFHLVLTTRGTIHKVV 440
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 461 VLPNNVSVELLELEVEFKNHAPITTKISSKQOLYVSSNEGYSOVLHRCIHYGA 520
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 441 E-PGEQHSFAFNIMEIQPRRAAIIOTMSIDARRLIYSSQWESQVPLDCEYVGGG 499
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 521 CADCLARPYCAMDGHSGSREYPTGKRSRRODVNRHGNPLQCGGFNKAIRNAAEIYQ 580
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 500 CHGCLMSRDPYCGWDGRCISITYSSE--RSVLQSLNPAEPHKECP--NPKPAFLQKVS 555
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 581 YGVKNNTFELCAPKSPQASIKWLQKDK 609
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 556 LAP-NSRYLSCPMESRAHATYSWR-HKEN 582
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 9 PRELIMINARY: PRT: 751 AA.
ID 042236 AC 042236:
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
```

```
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE COLLARSIN 3
OC GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RA FEINER L., KOPPEL A.M., KOBAYASHI H., RAPER J.A.;
RL NEURON 19:0(1997).
DR EMBL: AF022946; G2522204; -.
SQ SEQUENCE 751 AA; 85432 MW; 0B9BA2F7 CRC32;

Query Match 14.2%; Score 705; DB 13; Length 751;
Best Local Similarity 33.5%; Pred. No. 2,19e-142; Indels 41; Gaps 28;
Matches 150; Conservative 107; Mismatches 150;

Db 167 SENPNVNTSVYVINEELFSGMYI-DFMGTDALFSLTKRNAVVRTDQHSKWLSEPIFVD 225
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 171 PFSPEPNSLVLFEGDEYVSTIRKQYNGK-IPFRFRIGSESELYTS--DTV-MQNPQFIK 226
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 226 AHVDPGTDPNDAKYFFFEKRLIDNDSGTQIHSMARICPNPTGGQRL-VNKNWTFEL 284
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 227 ATIVHOD-QAYDDKIYFFREDNPKNEAPLNSRVAQLCRGDGGESSLVSXWNTFL 285
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 285 KARLYCSVDEDEGPETHEDVELL-ETDNP-RTILVYGIFTTSSSVFKGSACVYHL 342
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 286 KAMLYCS--DA-ATKKNFNRLQDVFLPDPSCQWDRTRYGVFSNPMN-Y--SACVYSL 339
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 343 SDIQVFNGPFAHKEGPNHQLIYQGRIPYRPGTCPGGAFTPNMRTTEFPDDVYTFIR 402
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 340 GDIDVFRTS-SLK-G-----YHSSLPNRPKGLPQ-QP-IPTEF-F--QVAD-RHP 385
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 403 NHPLMYNSIYPIHKRPPLIRIGTDYKYTKIADVRYNADGRC-RYHFLFGTDRGVQKVV 461
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 386 EVAQREPMGFL-KTFLF-H--SKYHOKVAVHMQASHGETFHLVLTTRGTIHKVE 441
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 462 LPTNSASGELLELEVEFQNSPTTKISSKQOLYVSSNEGVTQVPLHRCRIYTAG 521
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 442 -PGEQHSFAFNIMEIQPRRAAIIOTMSIDARRLIYSSQWESQVPLDCEYVGGC 500
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 522 ADCCLARPYCAMDGHSGSREYPTGKRSRRODVNRHGNPLTQCGGFNKAIRNAAEIYQ 581
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 501 HCLMSRDPYCGWDGRCISITYSSE--RSVLQSLNPAEPHKECP--NPKPAFLQKVS 556
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 582 GYKNNTFLECTPKSPQASIKWLQKDN 609
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Oy 557 AP-NSRYLSCPMESRAHATYSWR-HKEN 582
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 10 PRELIMINARY: PRT: 753 AA.
ID 013372 AC 013372:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SEMAPHORIN III FAMILY HOMOLOG.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA TISUE-BRAIN.
RC MEDLINE: 96230324.
RX XIANG R.H., HENSEL C.H., GARCIA D.K., CARLSON H.C., KOK K., DALY M.C.,
RA KEBBACHER K., DEN BERG A., VELDHOIS P., BUIS C.H., NAYLOR S.L.;
RT "Isolation of the human semaphorin III/F gene (SEMA3f) at chromosome
17p32, a region deleted in lung cancer."
RL GENOMICS 32:39-48(1996).
DR EMBL: U38276; G1061351; -.
DR PFAM: PF00047; 19; 1.
SQ SEQUENCE 753 AA; 84941 MW; BCBEBB0 CRC32;
```

Query Match	13.8%	Score 688	DB 4	Length 753
Best Local Similarity	31.4%	Pred. No. 5,20e-18		
Matches 139	Conservative 121	Mismatches 140	Indels 42	Gaps 30
Db	171	PYDPKIDTASALINELVAGYII -PFMGTDAAIFRTLGKOTAMRTDQYNSFWLMDPSFIH	229	
Qy	171	PEPSPENSILVLEPGEDGVSTIRKQGYNNK -IPRFRRINGESELVTS --DIY-MONPOFIK	226	
Db	230	AEIIDSAMENDKLTFFFRERSAE -APOSPAYIAAIGICLINDOGHCL -VNRKSTFLK	287	
Qy	227	ATIVHODQAYDVKIYFFREDPNPNPAPINVSVAQLCGDGGESSLSVSKWTFELK	286	
Db	288	ARLVCSVPGEDDIETHPELODVEY -OOTODY -RNPVYIAFTSSGVFSSAVCYVMA	345	
Qy	287	AMLVCS --DAATKNKNFRLODVELLPDPSSQOMDIRYGVFSNP -WNY -SANCYVSLG	340	
Db	346	DIRMVFENPFPAKESPNQWMPFGSKMYP PRRPTCPGGSTTPPSKSTKYDVPDEVINMRS	405	
Qy	341	DIDKFERTS -SLK -G -YH ---SS -LNPSPGKCLPQD -QP -I -PIETF -GVAD --R	383	
Db	406	HPLMVQAVYPLQRPRLVVRT -GAPRLTITIVADOVDSADG -RYVELTGLDGRGVQKVIY	463	
Qy	384	HPVAVQARVEPMP -PLKTRPLFSKTHYOKAVAHKQASHGTEFVYLLTTRDGRITKHYVE	441	
Db	464	LPKDDQDEHELMLEVEVEYFKDPAPYKTIWTISSKQOOLVYASAVGVTLSLHRQAYGAAC	523	
Qy	442	-PGEDEHSAFENINEMIOFFRRAAAIQMTSLDAERKRKYLVSSOWEVSQVPLDCEYVGGC	500	
Db	524	ADCCIAARPYCAMPQOQACSFYTAASKRRSRRODVRRHNPTRCQGFNSMKNKNAVESQY	583	
Qy	501	HGCLMSRNPYCGMOGRIS -IYSSER -SVLQISINPAEPKRECF -NPKDKAPLOKVL	556	
Db	584	GVAGSAFLEOCPPSPQATYKW	605	
Qy	557	A -PNSRYVILSCPMSRSHATYSW	577	

[illegible]

```

QY 133 LERBSBG-LLACGTAARHPC-W-NL-VN-GIVVPL-G-EM-RCIAPFSPDENSIVL 181
Db 181 LVDSGLLYSGTA-ADFMGHDFAIFRTLGHHPDIRTEQHSRWMLNDPRISAHLPESDNPE 239
QY 182 FEGDVUUYTIKROENYK-IRFRIRIGESFLYNS--DTV-MQNPGIKATIVHQ-DQAY 236
Db 240 DDKIYFFERENAIQGEHTGKATTHARIOLOKNDGGRSL-VKWTTFKLALICGVPGR 298
QY 237 DDKIYFFEREDNPKNPRAPLANSRVACLOCGDGGGSSLSYVKMTFLKAMLYCS--DA 294
Db NGIDTHELODVFLM-NSKDP-KNPIYUGVFTTSSNIFKGSVACMSITDVARVFLGY 356
QY 295 -ATNNENRLODVLELDPDSGOMDTRVYGVSFPMVN-Y-SAVCYSLGIDDKVETRIS- 349
Db 357 AHRDSRPNQWPRYUGRVPYRPRGCSBKTEGGEDSTOLDPEVITLARSHPMYANVPFI 416
QY 350 SLK-G-YH--S--S-LPNRPGRCSLDRQ-OPi-PIETF--QY-A-DR-HPEAORVERBE 394
Db 417 NS-REIMIKTDVDOFTOIVDVRVADBDQ-YUMVEIGTFICVLVYVSLIPKETHLEBE 474
QY 395 GPKLPYRHSK-YHQYKAVNHRQASHGFTFHVLLTDRGIIKHVVE-PEQENSPAF 451
Db 475 VLEEMVYFRPPIYISAMKISTKOOLYIGSATGVSOBLPHRCOVYKACAECCSLARDPY 534
QY 452 NIME-IOPFRPAAALIQMSIDAERKLYVSSOWEVSQVPLDCEYVGGCGHCLSLSPDY 510
Db 535 CAMDGSSCSRFPPAKRRTRRQDIRNDPLTHGSDJONHNDPSGQLEKIIYGVNSST 594
QY 511 CGMDGRCRISITISE-NSVL-QSINAERPHKECPN-PRKDPKAP--LQ-KYSLA-PNSRY 562
Db 595 FLECSPKSQRAIY-WQFOKONDDHKYE 621
QY 563 YLSC-PMESRATYSWR-KHEENEGCE 588

```

[illegible]


```

RESULT 14 PRELIMINARY: PRT: 785 AA.
ID Q13275
AC Q13275; Q13274;
AC 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SEMAPHORIN IV.
OS HOMO SAPIENS (HUMAN).
OC EUAROTIA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
CC CATARRHINI; HOMINIDAE; HOMO.
LN 11
RN 1
RP SEQUENCE FROM N.A.
RA NELSON J., BERNALD T.
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN 12
RP SEQUENCE OF 394-436 FROM N.A.
RC TISSUE-PLACENTA:
RX MEDLINE; 96210603.
RA SEIDIO Y., BADER S., LATIF F., CHEN J.Y., DUH F.M., WEI M.H.,
RA ALANSI J.P., LEE C.C., LERAN M.I., MINNA J.D.;
RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
RT cancer deletion region and demonstrate distinct expression
RT patterns."
RL PROC. NATL. ACAD. SCI. U.S.A. 93:4120-4125(1996).
DR EMBL; AC000063; G1669373; -
DR EMBL; U32172; G995788; -
DR EMBL; U32171; G995786; -
PFAM; PF00047; 1g; 1
SQ SEQUENCE 785 AA; 88381 MW; CD175765 CRC32;
Query Match 13.7%; Score 681; DB 4; Length 785;
Best Local Similarity 31.4%; Pred. No. 3,27e-136;
Matches 139; Conservative 122; Mismatches 139; Indels 43; Gaps 31
DB 202 HYDRLDASALINEELVAGYVI-DEMGTDPAIRITLGKGTAMFTDQYNSRWLNDPSFIH 260

```

Qy 171 PPSDENSELVFEGDEVYSTIRKQENYNGK-IPFRRINGESELYTS--DTV-MONPOFIK 226

Db 261 AELLPSAERNDDKLYFFFRERSAE-APQSPAYVARIGRICLNDGDGHCL-VNKWSTFL 318

QY	227	ATIH0D-QAANDKXIYFFFRDNDNDKNDKPEAPLANSRAQJCRGQGGESSLSXKMMTFL	285
Db	319	KARLVCSYPGEGDGIETHEPDELQVYV-QQOTDV-RNPVIYAVFTSSGSVFRGSAVCYSM	376
QY	286	KAMLVCS---DAANKKFNENRLQDFELLDPDSSGGWRDRVYGVFSNP--WNY--SAVCYYSL	339
Db	377	ADIMWVFNFGPAHKEGNYQMFPFSGKMPYRPBETCGGFTPEBMSKTXYPDEVINEMR	436
QY	340	GDIDKVFRTS-SLK-G-YH---SS-LPNDRPKCLPDO-QP-I-PTKFF--QVAD--R	383
Db	437	SHPLMYOAVYPLQRRPRLVVRT-GAPYRLTIAVDQVDAADG-RXYELFELTDRGTQVKI	494
QY	384	-HPEVAQVREBMG--PLKTPLEHSHKYHQYKVAVHRMQASHGETFHVLYLTTRDTGIHKV	440
Db	495	VLPDDDELEBIMEEVEYFVRKDPAPVATMTIISKRQOLYASANGVTHLSLHROAAGAA	554
QY	441	EPGEQHSFA-FNMTLOPFRRAAIDMTSLDARRLYLWSSQWESQVPLDCEVYGGG	499
Db	555	CADCECLARDPYCAMPDGOACSRYSKRSRRRPOVRHGNLPIROCGRGNSNANKNAVESQ	614
QY	500	CHGGLMRDRPYCGMDQGRCS-YYSSSR-SVLQSLNAEERHKECP--NPRDKAPRLQKVS	555
Db	615	YGVAGSAFLECPQRPQATYVKM	637
QY	556	LA-PNSRYTLSCPMESHATYSW	577

DT 01-NOV-1998 (TREMBL:REL. 08, LAST ANNOTATION UPDATE)
DE SEMAPHORIN.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN.
RX MEDLINE; 96226360.
RA ROCHE J., BOLDOS F., ROBINSON M., ROBINSON L., VARELLA-GARCIA L.,
RA SWANTON M., WAGONER B., FISHEL R., FRANKLIN W., GEMMILL R.,
RA DRABKIN H.;
RT "Distinct 3p21.3 deletions in lung cancer and identification of a new
RT human semaphorin."
RL ONCOGENE 12:1289-1297(1996).
DR EMBL; U33920; G100207; -.
DR PFAM; PF00047; 1g; 1.
SQ SEQUENCE 785 AA; 88385 MW; 9971F14E CRC32;

Query Match 13.7%; Score 679; DB 4; Length 785;
Best Local Similarity 31.6%; Pred. No.1,07e-135;
Matches 140; Conservative 121; Mismatches 139; Indels 43; Gaps 31;

Db 202 PYDKPLDTASALINELYAGYI-DFMGTDALFRTLGQTAMRTDQYNSRWLNDPSFIH 260
QY 171 PFSPDENSLVLEGEDEVYSTRKOEYNGK-IPFRFRIGSESELYTS--DTV-MONPOFIX 226
Db 261 AELIPDSARNDKLYFFERSAE-AQSPAVYARIGRICLNDGHCCL-VNKWSTFL 318
QY 227 ATIVHOD-QAYDDKIYFFREDNPKNPEAPLNSVQAQLCRGDGSESSLSVSKWTF 285
Db 319 KARLYCSVGEDEGTEHFDLODFV-QOTQDV--RNPVIYAVFTSSGSVFRGSACVYSM 376
QY 286 KAMLYCS--DAATKNKNRLQDVFLPDPGQMRDTRVYGVFSNP-WNY--SAVCYISL 339
Db 377 ADIRAVENGPFARKEGPNYQMPFSGKMPYPRPGTCPGGFTFPKSKSTKDYPDEVINPMR 436
QY 340 GDIDKVFRTS-SLK-G-YH---SS-LPNRPKCLPDQ-QP-I-PTETP--QVAD--R 383
Db 437 SHPLMYQAVYPLQRRPLVVR-T-GAPYRLTTTAVDOVDAGDG-RYEVFLGTDRTGVOKVI 494
QY 384 -HPEVAQREVEPMG--PLKTPLEHSHKYHQAQVAVHRMQASHGETFHVLYLTDRGFIHKV 440
Db 495 VLPKDDQMEELMLEEVVFQDPAPVKTMTTSSKROQLYVASAVGVTHLSLHRCQAYGAA 554
QY 441 E-PGEQHSFANIMEIOPFRRAAIQTMSLDAERKLIYSSQWVEVSQVPLDLCEVYGGG 499
Db 555 CADCCCLARDPYCAMDGAQCSRYTASSKRRSRQDVRRGNPIROC GFNSNANKNAVESVQ 614
QY 500 CHGCLMSRDPCGMDQGRCS-IVSSER-SVLOSINPAEPHKCP--NPKPDKAPLQKVS 555
Db 615 YGVAGSAFLCOPRSPQATYKW 637
QY 556 LA-PNSRYLLSCPMESRHATYSW 577

Search completed: Tue Aug 3 16:32:25 1999
Job time : 127 secs.